Discussion of renewal process methods for estimating the incubation period of SARS-CoV-2: A simulation study

Leitzinger, N.

Mathematical Institute

Master Thesis

Statistical Science for the Life and Behavioural Sciences

Discussion of renewal process methods for estimating the incubation period of SARS-CoV-2: a simulation study

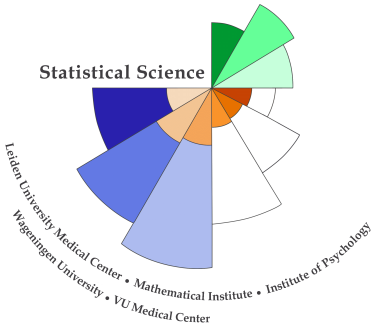
Author:
Nils Leitzinger

Supervisor:
Prof. Dr. Marta Fiocco
Mathematical Institute Leiden University &
Leiden University Medical Centre

External Supervisor:
Dr. Ronald Geskus
Oxford University Clinical Research Unit

November 2021

Universiteit Leiden
The Netherlands



Acknowledgments

First and foremost, I would like to thank my supervisors, Prof. Dr. Marta Fiocco and Dr. Ronald Geskus, for their guidance, continuous support, and patience. I am also grateful to Vera Artzen (MSc) for her support, which ranged from providing literature to checking my code. Furthermore, I would like to thank all the members of the OUCRU Covid-19 Modelling Group for allowing me to participate in their weekly discussions. Lastly, I would like to thank Prof. Dr. Hein Putter and Dr. Kanavetas for evaluating my work and taking part in the exam committee.

Abstract

One of the key characteristics to describe an infectious disease is incubation period. Commonly incubation period estimates are obtained via interval-censored methods. Deng, You, Liu, Qin, and Zhou (2020) and Qin et al. (2020) proposed a new family of methods for estimating incubation period and applied them to data from the initial SARS-CoV-2 outbreak in Wuhan. These methods are based on the theory of the renewal process and do not require information on the time of infection. Instead travel information (i.e., day of departure) are needed. These data tend to be easier obtainable and hence, larger datasets can usually be used. However, both Deng and Qin made a number of assumptions that appear questionable. To date, no study has addressed the validity of their proposed renewal methods nor their assumptions.

In a novel simulation study, the impact of changing assumptions on the estimated incubation time was investigated. Deng and Qin assumed that the time from infection to leaving Wuhan follows a uniform distribution. This assumption is problematic because of the exponential increase in SARS-CoV-2 cases and the sharp increase in people leaving Wuhan before lockdown measures were implemented. In addition, both assume that up to 20% additional infections occur at day of travel due to busy environments. However, it is not clear whether the correction for additional infections at travel day is warranted. As part of the thesis, a data generation method was introduced that takes these aspects into account.

In this thesis, it is shown that the assumptions underlying the renewal process method are violated by Qin and Deng and the proposed data generation method. The simulation study showed that the violated assumptions introduce a bias that is partially compensated by the bias introduced by the inclusion of additional infection at day of travel. The findings suggest that incubation period estimates based on current renewal process methods should be interpreted with caution. The results of this work provide important insights into the accuracy of current methods for estimating incubation period. This can help to better understand the dynamics of infectious diseases, which in turn can help to contain the spread of future outbreaks.

Contents

Chapter 1

Introduction

A brief introduction into SARS-CoV-2 can be found on Section 1.1. Infectious disease terminology relevant for the thesis are defined in Section 1.2. More detailed information on incubation period can be found in Section 1.4. Historical background on estimating infectious diseases characteristics is provided in Section 1.3, while the aims of the thesis and its structure are outlined in Section 1.5.

1.1 SARS-CoV-2

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by the severe acute respiratory syndrome – novel coronavirus 2 (SARS-CoV-2)(Wassie, Azene, Bantie, Dessie, & Aragaw, 2020). The most common symptoms include: fever, sore throat, headache, shortness of breath, and cough (Voysey et al., 2021). Many patients only show mild symptoms and it is estimated that around 50 % of patients are asymptomatic carriers (Voysey et al., 2021). COVID-19 first appeared in the Chinese city of Wuhan, Hubei province, in December 2019 and subsequently spread quickly around the world (Dong, Du, & Gardner, 2020). Four month after the virus was observed for the first time, the outbreak was declared a global pandemic and health emergency by the World Health Organization (WHO) (Ghebreyesus, 2020). Within the first year of the pandemic, approximately 2 million people worldwide died from COVID-19. Restrictions and lockdowns which were put in place to reduce the spread of the virus, triggered the largest global recession in history (World Bank, 2020). In the field of statistics, the COVID-19 pandemic resulted in a proliferation of research which focuses on mathematical modelling of infectious disease characteristics.

1.2 Terminology

Infectious diseases can be caused by a variety of *infectious agents/pathogens*. Infectious agents/-pathogens are defined as organisms that can produce an infectious disease (Krämer, Kretzschmar, & Krickeberg, 2010). They include parasites, bacteria, fungi, and viruses (Krämer et al., 2010). Transmission of pathogens and invasion of a *host* may occur via spread from person-to-person by direct contact, foodborne illness, water, insects, or via airborne aerosols (Krämer et al., 2010). Infectious diseases pose a major threat to human health. They account for a third of all deaths and in addition, are linked to a higher risk of developing serious ailments, such as cardiovascular diseases (Krämer et al., 2010). Various societal factors, such as urbanization, constant global transfer of goods, and high mobility that marks today's society, facilitate the spread of infectious diseases. Key characteristics to describe the spread of an infectious disease are *generation time*, *serial interval*, and *basic reproduction number*. The term *generation time* describes the time that passes between infections in an infector–infectee pair, whereas *serial interval* is defined

as the time between symptom onsets in an infector–infectee pair. Assuming a population in which all individuals are considered susceptible to an infection, the *basic reproduction number* R_0 is defined as the expected number of secondary cases infected by one primary case. A R_0 value larger than 1 indicates that the infection is able to spread in the population. Additional important quantities are *Incubation period* and *latent period*. Incubation period represents the time period that passes between infection and symptom onset, while latent period describes the time period that passes between infection and onset of infectiousness.

1.3 History of mathematical modelling of infectious diseases

The first attempt to express the spread of an infectious disease with the help of a mathematical model dates back to the 18th century (Dietz & Heesterbeek, 2000). In his work, Daniel Bernoulli introduced the concepts of force of infection and case fatality. He used both concepts to estimate the effect of smallpox vaccination on life expectancy (Dietz & Heesterbeek, 2000). However, it was not until the beginning of the 20th century that researchers were able to account for the non-linear properties that describe the spread of an infectious disease (Krämer et al., 2010). For AIDS, mathematical models were for the first time used to asses preventive measures and to predict the disease trajectory of the global epidemic. During the first SARS-CoV-1 outbreak, real time assessment of preventive measures was introduced. Over the course of the last few decades, mathematical models became crucial to understand and fight against the spread of infectious diseases. They play a particular important role in the early stages of an disease outbreak when pharmacological resources are scarce (Reich *et al.*, 2009). A key characteristic that drive these models is incubation period.

1.4 Incubation period

This thesis focuses on estimating a key characteristics of SARS-CoV-2: *incubation period*. Incubation period is defined as the time between infection and symptom onset. The virus load, the replication speed of pathogen after infection, and the defense mechanisms of the host appear to be the main factors that determine the length of incubation period (Wassie et al., 2020). Although the length of the incubation period can vary greatly from person to person, its distribution tends to be right skewed (Wassie et al., 2020).

In case of an outbreak, the incubation period is used to determine the length of quarantine for potentially exposed individuals. In this case, the focus is commonly on the upper bound of the incubation period. However, it should be noted that the latent period is theoretically better suited to determine the duration of quarantine. To break chains of infection, knowing the expected time from infection to the onset of infectiousness would be ideal (i.e., latent period). Since infectiousness is more difficult to monitor, daily testing would be required, incubation period is commonly used as a proxy of latent period.

In addition, the length of the incubation period can also provide information on the effectiveness of certain disease control measures. For example, when the goal is to contain the spread of an infectious disease with a relative long incubation period, isolation measures are known to be less effective (Nishiura, 2007). The main challenge in estimating the incubation period is incomplete data. This holds especially true at an early stage of an infectious disease outbreak when contract-tracing data are difficult to obtain (Deng et al., 2020). Various methods have been proposed to deal with the uncertainty surrounding the time differences between infection - symptom onset; most commonly interval-censored methods have been used. In the first few months of the COVID-19 pandemic, new approaches have been introduced that are based on the theory of the renewal process.

1.5 Aims and structure of the thesis

Qin et al. (2020) introduced the idea of using the theory of the renewal process to estimate incubation periods. Qin's method was subsequently refined by Deng et al. (2020). Both applied their method to data from the initial SARS-CoV-2 outbreak in Wuhan. To the best of the authors knowledge, no study has so far looked at the validity of the proposed renewal process approach. This thesis discusses Qin's approach and provides a detailed discussion of Deng's method. Deng's and Qin's methods offer advantageous over existing methods. For example, whereas standard approaches require information on the date of infection and symptom onset, the renewal process approaches require only the date of symptom onset and travel information. This is especially advantageous during an early phase of an outbreak when contact-tracing data are particularly difficult to obtain. This thesis focuses on renewal process methods and its underlying assumptions. Limitations of the methods are discussed.

For example, Deng and Qin both assume that the time from infection to leaving Wuhan is uniformly distributed. This assumption is problematic given that the number of cases increased exponentially during the SARS-CoV-2 outbreak in Wuhan. This thesis presents a data generation method that mimics the outbreak in Wuhan, taking into account the exponential increase in infections. Based on this data generation method, estimates of the incubation period are obtained using a novel simulation study. The estimates will be compared with Deng's, Qin's, as well as with estimates based on interval-censored methods.

The results of this work provide important insights into the accuracy of current methods for estimating incubation period. This may help to better understand the dynamics of infectious diseases, which in turn may help to limit the spread in future outbreaks.

The thesis is outlined as follows: Chapter 2 presents various methods for estimating the incubation periods, including Qin's, Chapter 3 focuses on Deng's approach and Chapter 4 provides a critical review of Deng's handling of the interval-censored approach. All details concerning the novel simulation study set up and results are discussed in Chapter 5. The thesis ends with a discussion, which can be found in Chapter 6.

Chapter 2

Estimation of incubation period & generation time

Several methods have been proposed to estimate the incubation period (Section 2.1). This chapter provides a literature review on the topic of incubation period (Section 2.2). As part of Section 2.2, a method proposed by Qin et al. (2020) for estimating incubation period based on a renewal process is outlined.

2.1 Incubation period

The incubation period, denoted here as I, can be computed as $I = S - E$, with S representing the time of symptom onset and E the time of exposure to the infection (Reich, Lessler, Cummings, & Brookmeyer, 2009). Let $f_\theta(i)$, dependent on a parameter θ, be the probability density function (PDF) of the continuous non-negative random variable I. The density can be expressed as

$$f_\theta(S - E) = f_\theta(I). \tag{2.1}$$

In practice, the time of infection is rarely known; hence, the time of exposure is commonly said to fall in a certain interval. When time of infections is not exactly observed, let the single interval-censored be defined as $(I_L, I_R) = (S - E_R, S - E_L)$ (Reich et al., 2009). I_L and I_R denote the incubation period with E_R and E_L being the right and left end points of the censored exposure interval. The potential role of the exposure interval as part of the single censored-interval likelihood function is described by Groeneboom (2021). He defined the log likelihood of the single-interval censored approach as

$$log \int_{t \in [0, E_i]} g(S_i - t) dF_i(t), \tag{2.2}$$

where the time of infection t is known to fall in the interval $[0, E_i]$ and F_i marks the cumulative distribution function (CDF) of the exposure time. It is commonly assumed that the exposure time E is uniformly distributed on the interval $[0, E_i]$. When maximizing the log-likelihood, the uniform distributed exposure time E does not play a role. Hence, we obtain

$$\begin{aligned} &\sum_{i=1}^{n} log\{ \int_{t=0}^{E_i} g(S_i - t) dt / E_i \} \\ =&\sum_{i=1}^{n} log\{ \int_{t=0}^{E_i} g(S_i - t) dt. \end{aligned} \tag{2.3}$$

Continuing the previous notation, the resulting likelihood can therefore be denoted as

$$L(\theta; I_L, I_R, S) = \int_{E_L}^{E_R} f_\theta(S - E)\ de. \tag{2.4}$$

If infection time and symptom onset are both not observed but instead are only known to lie within an interval, an observation is considered doubly interval-censored (Reich et al., 2009). An observation in this context is defined as $X = (E_L, E_R, S_L, S_R)$ where the subscripts L and R represent the left and right boundaries on the possible infection and symptom onset times. Assuming that infection and incubation time are independent, the likelihood for a doubly interval-censored observation is

$$L(\theta, \lambda; X) = \int_{E_L}^{E_R} \int_{S_L}^{S_R} g_\lambda(e) f_\theta(s - e) \, ds \, de \tag{2.5}$$

where $g_\lambda(e)$, dependent on the parameter λ, is defined as the PDF of the exposure time. Commonly research has used parametric models, e.g., gamma, Weibull, and log-normal, to describe the PDF of the incubation period (e.g., Backer, Klinkenberg, and Wallinga (2020); Reich et al. (2009)).

2.2 Estimation of SARS-CoV-2 incubation period

To date, only a limited number of studies have been published that address the estimation of the SARS-CoV-2 incubation period.

Lauer et al. (2020) examined Chinese news articles and public health records in order to gather information on SARS-CoV-2 cases during the initial phase of the outbreak in Wuhan. Lauer used the likelihood expressed in (2.5). They assumed that the PDF of the exposure interval follows a uniform distribution and the PDF of incubation period follows a log-normal distribution. The majority of people in their sample are either Wuhan residents or travelers who were infected during their stay in Wuhan. However, the sample also includes 20 people who were infected outside of Wuhan. For these people, the infector is known. The total sample size is 181. Lauer solely included cases in their sample for which the symptom onset occurred outside Hubei province. At that time, hardly any cases were officially known outside the province. In this way, the exposure interval for the people who had travelled to Wuhan could be determined based on the arrival and departure time. For all the other people included in the sample, the upper bound of the exposure interval was marked by the latest possible time of symptom onset, whereas the lower bound of the exposure interval was either defined by: 1) the earliest possible date of exposure to an infectious person, 2) 1 December 2019 when the infection was of unknown origin.

Backer et al. (2020) worked in the Bayesian framework to estimate the incubation period of SARS-CoV-2. They included 88 people who developed symptoms after leaving Wuhan. Both, Wuhan residents and people who visited Wuhan were comprised in the sample. While the time of symptom onset is known, the time of infection was not observed. Backer assumed a uniform prior probability distribution for the exposure intervals and used a single interval-censored likelihood (2.4). The authors fitted three parametric distributions for the PDF of the incubation period: Weibull, gamma, and lognormal. In combination with the known arrival and/or departure times from Wuhan, the authors inferred the date of exposure. This method has been criticised for introducing two type of biases:

1. A short follow-up period after leaving Wuhan, which reduces the chances of people with longer incubation periods to be included,

2. Because only people who developed symptoms after leaving Wuhan were included in the sample, people with short incubation period were less likely to be included (Qin et al., 2020).

Linton et al. (2020) used Bayesian methods in combination with the doubly interval-censored likelihood function (2.5) to estimate incubation period. Like Backer, Linton included Wuhan

residents and people who traveled to Wuhan. Again, the location of the boundaries of each interval was determined by the arrival/departure time to/from Wuhan. However, the author included an analysis that excludes Wuhan residents. It is argued that this leads to more precise estimates, as usually people who travelled to Wuhan have a more finely defined exposure interval due to their date of arrival. They assumed that the PDF of incubation period follows a Weibull, gamma, and log-normal distribution. The PDF of exposure was modelled as a uniform distribution. To overcome bias number 1. from Backer's study, correction for right truncation is taken into account via

$$f'(s-e,e) = \frac{f(s-e)}{\int_0^{T-e} \frac{r\ \exp(-ru)}{1-\exp(-ru)} F(T-e-u)du}, \tag{2.6}$$

where $f(\cdot)$ is the PDF of the incubation period, s is the time of symptom onset, e is the time of exposure to the infection, r is the assumed exponential growth rate (i.e., 0.14 (Jung et al., 2020)), T the last day of observation, and $F(\cdot)$ the CDF of $f(\cdot)$. The incubation period estimates obtained by Lauer and Backer can be found in tale 2.1. The estimates obtained by Linton when Wuhan residents are excluded from the analysis and right truncation is applied, are also shown in Table 2.1.

As noted by Qin et al. (2020), correcting for right truncation does not address bias number 2. To overcome both biases, Qin proposed an alternative approach based on a renewal process. A renewal process is defined as a sequence $\{X_i, i = 1, 2, \ldots\}$ of non-negative, independent, and identically distributed random variables (Cox, 1964). Let

$$T_n = \sum_{i=1}^{n} X_i \tag{2.7}$$

describe the point in time when the n-th renewal is taking place and $N(t) = max\{n : T_n < t\}$ the number of renewals in (0, t]. He defined the forward time, the time between any given time t and the next renewal, as

$$V(t) = T_{N(t)+1} - t, t > 0, \tag{2.8}$$

whereas the time the last renewal occurred (i.e., backward time) is given as

$$A(t) = t - T_{N(t)}. \tag{2.9}$$

Qin made use of the fact that at an early stage of the disease outbreak, a large number of cases left Wuhan asymptomatic and developed symptoms after leaving the city. Qin split the time from infection to symptom onset into backward time A and forward time V. The backward time is defined as the time from infection to departing Wuhan and forward time as the from departure to symptom onset. While the forward time is observed, the backward time is not known. For a visualization of backward and forward time, see Figure 2.1.

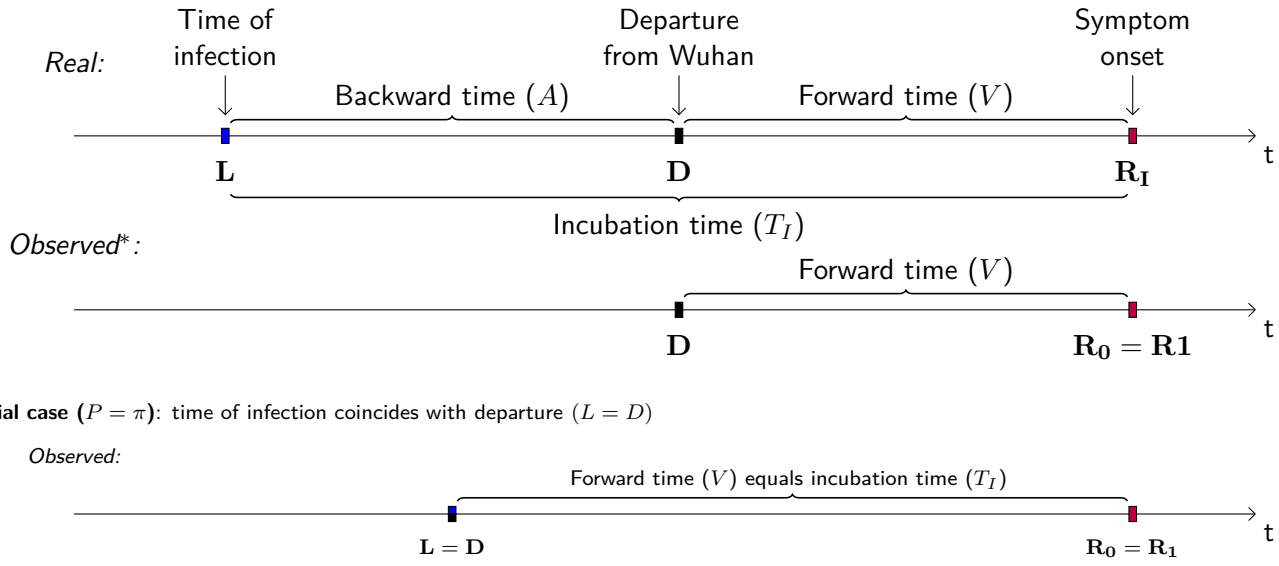


Figure 2.1: Visualization of backward and forward time.

Let $f(y)$ be the probability density function of the incubation period Y. The marginal distribution of A and V is given as

$$g(v) = \frac{\bar{F}(v)}{\mu}, \quad v \geq 0, \tag{2.10}$$

where $\bar{F}(\cdot)$ is the survival function. Qin assumes a uniform density on the time from infection to departure. Furthermore, it is assumed the incubation period follows a Weibull distribution with shape and scale parameters α and λ

$$f(y) = \alpha\lambda(y\lambda)^{\alpha-1}\exp\{-(y\lambda)^{\alpha}\}, \quad y \geq 0. \tag{2.11}$$

where Γ denotes the gamma function. The survival function of the Weibull distribution is $\exp\{-(\lambda v)^{\alpha}\}$ and its mean is $\frac{\Gamma(1+1/\alpha)}{\lambda}$. By combining (2.10)-(2.11) the density function for the forward time when a Weibull distribution is used can therefore be expressed as

$$g(v) = \alpha\lambda\frac{\exp\{-(\lambda v)^{\alpha}\}}{\Gamma(1/\alpha)}, \quad v \geq 0. \tag{2.12}$$

The detailed computations from (2.10) to (2.12) can be found in appendix 8.1.1. The likelihood to estimate the parameters of the Weibull is as follows

$$L(\alpha, \lambda) = \prod_{i=1}^{I} \alpha\lambda\frac{\exp\{-(\lambda v_i)^{\alpha}\}}{\Gamma(1/\alpha)}, \quad \alpha > 0, \lambda > 0, \tag{2.13}$$

where v_i represents the observed forward times. Since, Qin assumes that the crowded environments on the day of travel elevate the risk of infection, some forward times are incubation periods. The observed time from departure to symptom onset are hence, a mixture of the distributions of incubation periods and forward times

$$h(v) = (1-\pi)g(v) + \pi f(v) \tag{2.14}$$

where π expresses the percentage of people who were infected at day of travel. The Weibull distribution (2.14) becomes

$$h(v) = \alpha\lambda\{\pi(v\lambda)^{\alpha-1} + (1-\pi)/\Gamma(1/\alpha)\}\exp(-(v\lambda)^{\alpha}), \qquad v \geq 0. \tag{2.15}$$

For the detailed steps from (2.14) to (2.15), see 8.1.2 in the appendix. Note that without additional information the mixture model (2.14) is not identifiable nonparametrically; also in some parametric cases (like the exponential distribution), π is not identifiable (Qin et al., 2020). However, if $\alpha \neq 1$, model (2.15) is identifiable. As π is not known, it is estimated to take the following values: 0, 0.05, 0.1, and 0.2 to investigate the sensitivity of estimates of incubation period. The parameters α and λ are estimated by maximizing the following likelihood function

$$L(\alpha, \lambda) = \prod_{i=1}^{I} \alpha\lambda\big\{\pi(v_i\lambda)^{\alpha-1} + (1-\pi)/\Gamma(1/\alpha)\big\} \exp(-(v_i\lambda)^{\alpha})\big\}. \tag{2.16}$$

The estimated median, mean, and 95% percentile obtained by Qin are shown in Table 2.1. The table also includes estimates by Deng et al. (2020). For a discussion of Deng's method and results see Chapter 3.

Table 2.1: Estimates for incubation period. The unit of time for all estimates are days.

Study	Distribution	Median (95%CI)	Mean (95%CI)	95% percentile (95%CI)
Backer	Weibull	6.4 (5.5-7.5)	6.4 (5.6-7.7)	10.3 (8.6-14.1)
	Gamma	6.1 (5.3-7.3)	6.5 (5.6-7.9)	11.3 (9.1-15.7)
	Log-normal	6.1 (5.2-7.4)	6.8 (5.7-8.8)	13.3 (9.9-20.5)
Lauer	Log-normal	5.1 (4.5-5.8)	-	10.06 (7.5-13.2)
Linton	Gamma	5.2 (4.1-7.0)	-	13.4 (10.0-21.8)
	Log-normal	4.6 (3.7-5.7)	-	12.3 (9.1-19.8)
Qin	Weibull	8.13 (7.37–8.91)	-	16.67 (15.94-17.32)
Deng	Weibull	8.10 (6.40-8.67)	8.60 (7.03-9.08)	-
	Gamma	8.50 (7.22-9.15)	9.10 (7.86-9.66)	-
	Log-normal	8.74 (8.02-9.36)	9.44 (8.81-9.99)	-

Chapter 3

Estimation of incubation period and generation time based on Deng et al. (2020)

This chapter provides a summary of the paper by Deng and colleagues titled *Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China*. Section 3.1 contains information on the study design and the data selection. Section 3.2 outline Deng's approach to estimating incubation period. Section 3.3 entails information on the data generation method, the settings of the simulation study, and its results.

3.1 Introduction

Deng's renewal process approach to estimating incubation period of SARS-CoV-2 is closely related to Qin's (Section 2.2); both approaches only require dates of symptom onset and travel history. In their sample, Deng solely included people who left Wuhan between the 19^{th} and 23^{rd} of January 2020 and developed symptoms afterwards. Their dataset comprises a total of 1211 people. For each person, the time between leaving Wuhan and symptom onset is recorded (i.e., forward time; see Figure 2.1). This is the same dataset and dates as used by Qin. According to Qin, the severity of the virus was not widely acknowledge prior to the 19th of January. From the 19^{th} onwards testing and strict containment measures were started to get implemented throughout China. Thus, selecting the 19^{th} as the start of the sample window reduced the risk that people who left Wuhan infected each other after leaving the city (Qin et al., 2020). Based on previous studies on the incubation period of SARS-CoV-2, a maximum incubation period of 25 days was assumed (Qin et al., 2020). Since the 23^{rd} of January was the last day people could depart from Wuhan, February 15 marked the end of follow-up.

Like Qin, Deng assumes that the crowded environments at day of travel pose a elevated risk of infection. The percentage of people who got infected at day of travel (i.e., the complete incubation period) is represented by π (Deng et al., 2020). Hence, Deng models the observed data as a mixture distribution which consists of:

- the forward time
- the complete incubation period.

It should be noted that even so π is included, the sample is considered to be length-biased as people who developed symptoms before leaving Wuhan were excluded, people with shorter incubation periods were less likely to be included in the sample. The renewal process approach allows us to obtain estimates from a data type that has not been possible before. Given that

data regarding travel history tend to be easier to collect than dates of infection, sample sizes tend to be larger. This is results in more robust estimates (Deng et al., 2020).

3.2 Deng's approach to estimate incubation period of SARS-CoV-2

Similar to Qin's PDF of forward time, Deng's PDF forward time $h(\cdot)$ is given by the function $S(\cdot)$ divided by the expected value of the incubation period I

$$h(t) = \frac{S(t)}{E(I)} = \frac{\int_t^{+\infty} f_I(y)dy}{\int_0^{+\infty} y f_I(y)dy}, \quad t > 0, \tag{3.1}$$

where $f_I(\cdot)$ represents the PDF of the incubation period. Again a uniform distribution of time from infection to departure is assumed. Identically to Qin, only the forward times $t_1, ..., t_m$ for individuals $j = 1, .., m$ were observed; not the incubation periods. Hence, the incubation period I used in equation (3.1) is not observed. Furthermore, based on the observed forward time it is not known who got infected at the day of travel and who before. Therefore, a mixture distribution is used where π denotes the unknown probability of getting infected at day of departure, $1-\pi$ the unknown probability of getting infected prior to the day of departure. Furthermore, θ represents the model parameter for the PDF of incubation period $f_I(\cdot)$ and the forward time $h(\cdot)$

$$Q(t; \theta, \pi) = \pi f_I(t; \theta) + (1 - \pi)h(t; \theta), \quad t > 0. \tag{3.2}$$

Details regarding the likelihood of the mixture distribution can be found in appendix 8.2.1. When accounting for errors in daily reports by assuming censored intervals, $t_j^+ = t_j + 0.5$ and $t_j^- = t_j - 0.5$, the estimates for π and θ can be obtained by maximized the following likelihood function:

$$L(\theta, \pi; t_1, ..., t_m) = \prod_{j=1}^{m} \Big[\pi\big\{F_I(t_j^+; \theta) - F_I(t_j^-; \theta)\big\} + (1 - \pi)\big\{H(t_j^+; \theta) - H(t_j^-; \theta)\big\}\Big], \tag{3.3}$$

where F_I and H represent the cumulative distribution function (CDF) of the incubation period I and the forward time V in the renewal process. In order to be able to model the forward time as a renewal, a number of assumptions need to be made:

1. The renewal process has reached the equilibrium state.
2. Individuals included in the sample were indeed infected in Wuhan and developed symptoms after departure.
3. The incubation periods follows a continuous distribution.
4. The first moment of the incubation period distribution exists and is finite.
5. The incubation periods are independent and identically distributed (i.e, IID).

In the long run a renewal process reaches its equilibrium status (Qin et al., 2020). Given that millions of people left Wuhan before the lockdown started on 23^{rd}, Deng claims that assumption 1. is very likely satisfied. The equilibrium status indicates that the renewal process has become reversible. Consequently, the data that is modelled by the renewal process and its time-reversed data, share the same distribution (Qin et al., 2020). When mirroring the unit of measurement on the x-axis (i.e., day x to departure and departure to day x) the distributions of the forward and backward can hence be seen as exchangeable. Assumption 2 is likely to be satisfied since at the beginning of the pandemic the risk of contracting the disease outside Wuhan was very low (Qin et al., 2020). The same holds true for assumption 3 and 4. Since the incubation period can take on any value between 0 and any finite number, the incubation period can be assumed to be continuous. Assumption 5 is a standard assumption.

3.3 Set up of Deng's simulation study & results

In Deng's simulation study, the incubation period is assumed to follow three distributions: gamma distribution ($k = 5, \theta = 0.8$), Weibull distribution ($\lambda = 2, k = 8$), and log-normal distribution ($\mu = 1.8, \sigma^2 = 0.4^2$). Three option for sample size (600, 1200, and 1800) and three values for π (0, 0.01, 0.2) were tested. Each setting was repeated 1000 times. Deng's approach (3.3) is contrasted with Qin's (2.16) and an interval-censored method. Deng's interval censored method and its problematic data generation approach is discussed in Chapter 4. Deng's data generation method is described in Algorithm 1.

Algorithm 1 Deng's data generation method (excluding the IC-method)

```
for (run in 1:1000) do
    for (i in 1:sample size) do
        with probability π:
        x <- draw a complete incubation period
        with probability 1-π:
        C <- value ranging from 1 to 30 from uniform distribution
        Y <- sample complete incubation period
        if Y > C then
            x <- Y - C
```

This algorithm mimics the distinction between infections prior to leaving Wuhan (probability 1-π) and at day of travel (probability π). x corresponds to a "complete" incubation period. For these people the forward time and incubation period have the same length, as they were infected at day of departure. The backward time C repents the time from infection to leaving Wuhan. Further, as specified by Deng, people who developed symptoms before leaving Wuhan were not included in the sample ($Y > C$). For those who developed symptoms after departing from Wuhan, Deng computes the forward time x (Y-C). Table 3.1 provides an overview of the simulation study results. Estimates and 95% confidence intervals are provided for $\pi = 0, 0.2$ and a sample size of 1800. The results for all simulation studies based on Algorithm 1 can be found in Appendix 8.3.4. Deng incubation period estimates are slightly larger than the published estimates that did not use a renewal process approach (see table 2.1). It is argued that this due to the longer follow-up period used by Deng. The small deviation to Qin is explained by the introduction of the interval censored daily reports. According to Deng, this results in more robust estimates.

Table 3.1: Simulation study based on Algorithm 1

Gamma		With $\pi = 0.0$		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\alpha}$	5	4.53 (4.48-4.58)	4.52 (4.47-4.57)	5.28 (5.21-5.34)	5.27 (5.20-5.34)
$\hat{\beta}$	0.8	0.76 (0.76-0.77)	0.76 (0.76-0.77)	0.82 (0.81-0.83)	0.82 (0.81-0.82)
$\hat{\pi}$		0.08 (0.07-0.08)	0.07 (0.07-0.08)	0.19 (0.18-0.19)	0.18 (0.18-0.19)
Mean	6.25	5.89 (5.86-5.92)	5.89 (5.86-5.92)	6.37 (6.34-6.40)	6.37 (6.34-6.40)
0.25Q	4.21	3.86 (3.83-3.88)	3.86 (3.83-3.89)	4.33 (4.30-4.37)	4.34 (4.31-4.38)
Median	5.84	5.45 (5.42-5.48)	5.46 (5.43-5.49)	5.96 (5.93-6.00)	5.98 (5.94-6.01)
0.75Q	7.84	7.45 (7.42-7.48)	7.46 (7.43-7.49)	7.97 (7.93-8.00)	7.99 (7.95-8.02)
0.90Q	9.99	9.61 (9.59-9.64)	9.63 (9.60-9.66)	10.11 (10.08-10.14)	10.14 (10.10-10.17)
0.95Q	11.44	11.08 (11.06-11.11)	11.10 (11.07-11.13)	11.56 (11.53-11.59)	11.59 (11.56-11.62)
0.99Q	14.51	14.21 (14.18-14.24)	14.23 (14.20-14.26)	14.61 (14.58-14.64)	14.65 (14.62-14.68)
MSE $\hat{\alpha}$		0.83	0.80	1.30	1.33
MSE $\hat{\beta}$		0.01	0.01	0.01	0.01
MSE $\hat{\pi}$		0.01	0.01	0.01	0.01
Weibull		With $\pi = 0.0$		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\lambda}$	2	1.93 (1.92-1.93)	2.00 (1.99-2.01)	2.04 (2.03-2.06)	2.18 (2.17-2.19)
$\hat{k}$	8	7.59 (7.56-7.63)	7.97 (7.96-7.99)	8.17 (8.12-8.22)	8.74 (8.71-8.77)
$\hat{\pi}$		0.08 (0.07-0.08)	0.01 (0.00-0.01)	0.19 (0.18-0.19)	0.09 (0.08-0.09)
Mean	7.09	6.74 (6.71-6.78)	6.74 (6.71-6.78)	7.24 (7.20-7.29)	7.24 (7.20-7.29)
0.25Q	4.29	3.98 (3.95-4.01)	4.28 (4.26-4.29)	4.45 (4.40-4.49)	4.94 (4.91-4.96)
Median	6.66	6.28 (6.24-6.32)	6.64 (6.62-6.65)	6.83 (6.78-6.87)	7.39 (7.36-7.42)
0.75Q	9.42	9.00 (8.96-9.04)	9.39 (9.37-9.41)	9.58 (9.54-9.63)	10.16 (10.12-10.19)
0.90Q	12.14	11.72 (11.68-11.76)	12.11 (12.09-12.12)	12.29 (12.25-12.34)	12.82 (12.79-12.86)
0.95Q	13.85	13.45 (13.41-13.48)	13.82 (13.80-13.83)	13.99 (13.95-14.03)	14.47 (14.44-14.51)
0.99Q	17.17	16.83 (16.79-16.87)	17.14 (17.12-17.17)	17.29 (17.25-17.33)	17.64 (17.60-17.68)
MSE $\hat{\lambda}$		0.03	0.01	0.04	0.05
MSE $\hat{k}$		0.56	0.07	0.64	0.83
MSE $\hat{\pi}$		0.02	0.00	0.02	0.02
Log-normal		With $\pi = 0.0$		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\mu}$	1.8	1.76 (1.75-1.76)	1.76 (1.76-1.76)	1.81 (1.80-1.81)	1.81 (1.81-1.82)
$\hat{\sigma}$	0.4	0.41 (0.41-0.41)	0.41 (0.41-0.41)	0.40 (0.40-0.40)	0.40 (0.39-0.40)
$\hat{\pi}$		0.05 (0.04-0.05)	0.04 (0.04-0.05)	0.19 (0.19-0.20)	0.19 (0.18-0.19)
Mean	6.25	6.31 (6.29-6.34)	6.31 (6.29-6.34)	6.61 (6.59-6.63)	6.61 (6.59-6.63)
0.25Q	4.21	4.39 (4.37-4.41)	4.42 (4.40-4.44)	4.68 (4.66-4.70)	4.71 (4.69-4.73)
Median	5.84	5.80 (5.77-5.82)	5.83 (5.81-5.85)	6.11 (6.09-6.13)	6.14 (6.12-6.17)
0.75Q	7.84	7.66 (7.63-7.68)	7.69 (7.67-7.71)	7.98 (7.96-8.00)	8.02 (8.00-8.05)
0.90Q	9.99	9.83 (9.81-9.86)	9.87 (9.85-9.90)	10.16 (10.13-10.18)	10.20 (10.18-10.22)
0.95Q	11.44	11.42 (11.40-11.45)	11.47 (11.44-11.49)	11.73 (11.71-11.75)	11.78 (11.76-11.80)
0.99Q	14.51	15.14 (15.11-15.17)	15.18 (15.16-15.21)	15.38 (15.35-15.41)	15.43 (15.40-15.46)
MSE $\hat{\mu}$		0.01	0.01	0.00	0.00
MSE $\hat{\sigma}$		0.00	0.00	0.00	0.00
MSE $\hat{\pi}$		0.01	0.01	0.00	0.00

Chapter 4

Critical appraisal of Deng's interval censored data generation method

Deng et al. (2020) contrasted his method (3.3) with Qin's method (2.16) and an interval-censored approach. Deng uses the interval-censored method proposed by Backer et al. (2020) and Linton et al. (2020) (see section 2.2). However, due to erroneous data generation for the latter approach, Deng's interval-censored method returns incorrect estimates. Therefore, Deng's interval censored results have not been included in Section 3.1 and this approach is instead separately discussed in this chapter. The data generation method for Deng's and Qin's method are shown in Algorithm 2. The parts only necessary for the interval-censored method, not included Algorithm 1, are marked in red. This algorithm again distinguishes between the probability of getting infected at day of travel (π) and prior to departure from Wuhan (1-π). Line 5 represents the start of exposure interval c for persons infected on the day of travel and line 11 for persons infected before the day of departure.

Algorithm 2 Deng's data generation method

```
1:  for (run in 1:1000) do
2:      for (i in 1:sample size) do
3:          with probability π:
4:          x <- draw a complete incubation period
5:          c <- round a randomly drawn value from uniform distribution [0,30]
6:          with probability 1-π:
7:          C <- drawn value from uniform distribution [0,30]
8:          Y <- sample complete incubation period
9:          if Y > C then
10:             x <- round(Y - C)
11:             c <- round a randomly drawn value from uniform distribution [0,30]
```

The distributions of the incubation period Y and x in line 4 correspond to the distributions specified in the simulation settings. The sample size options are again 600, 1200, and 1800 and 0 and 0.2 for π. The backward time C still repents the time from infection to leaving Wuhan. For those who developed symptoms after departure, Deng computes the forward time (Y-C) and samples a value between 0 and 30 for each forward time from a uniform distribution. The latter value indicates the start of the exposure interval c. In this way, Deng attempts to artificially create interval-censored data. However, his approach is problematic.

Lets assume $C = 8$, $c = 5$, and $Y = 10$. This results in the backward time to be 8 and the forward time x to be 2, even so the exposure interval has is only of length 5. Hence, by resampling the start of the exposure interval, there is the possibility that the exposure interval will not contain the infection time. The R-code used by Deng to generate the data is available

on GitHub `https://github.com/naiiife/wuhan` under incubation S1-S3.

Chapter 5

Novel simulation study

This chapter introduces a novel simulation study to estimating incubation period of SARS-CoV-2. Section 5.1 provides an overview how the outbreak in Wuhan was mimicked. Information on the simulation settings and the results can are shown in Sections 5.2 and 5.3.

5.1 Adapted data generation method

Deng et al. (2020) assumed a uniform distribution for the time from infection to leaving Wuhan. This assumption is problematic given the exponential increase of cases at the beginning of the Wuhan outbreak (Zhao et al., 2020). This simulation study investigates the usefulness of Deng's renewal method approach when an infectious disease spreads exponentially. A new data generation method is proposed to simulate the growth in cases in Wuhan between the 5^{th} to 23^{rd} of January 2020 (details are provided in Algorithm 3). This method accounts for the exponential growth in new cases. In addition, it takes the increase in people leaving the city due to looming lockdown measures in Wuhan and Chinese new year into account (which was on the 25^{th} of January). With the help of `for-loop`, each day is simulated individually. Each day the current number of cases is being multiplied with the growth rate, starting with 125 positive SARS-CoV-2 cases on January 5. To select the newly infected individuals for a given day, individuals are drawn randomly from the total population of Wuhan (11 million) in the amount of new cases. Subsequently, as mentioned above, each day the number of cases increases based on a fixed growth rate by drawing newly infected from the a total population of Wuhan. For the first 15 days of this 20 days window, it was assumed that 245.000 people left and entered the city of Wuhan each day. From the 19^{th} to the 23^{rd}, it was assumed that the number of people leaving Wuhan daily increased to 490.000. Further, it was assumed that over the course of the last five days before the lockdown was implemented, no people from outside Wuhan had entered the city. The number of people entering and leaving the city was chosen so that the sample size per simulation varied around the 1211 cases found by Deng. From the 19^{th} onwards, for people who left Wuhan infected, an incubation period is drawn from one of the parametric distributions used by Deng (as outlined in Section 3.3). However, only those infected travelers who have not experienced symptoms when leaving Wuhan are stored. For those, the forward time is returned as the sum of infection time and incubation period minus the day `i` as specified by the `for-loop`. See Histogram 5.1 for a comparison between the forward and backward times based on Algorithm 1 and 3 and Deng's and Qin's data generation method.

Algorithm 3 Simulate Wuhan Outbreak

```
Total cases <- 125
for i in 1:19 (represents 5th to 23th of January) do
   List of infected[[i]] <-random sample without replacement from total
population of size of Total cases
   if i < 15 (before 19th of January) then
      Add 150.000 people to total population
      Sample without replacement 150.000 people from total population
      Remove sampled 150.000 people from total population
   end if
   if i >= 15 (19th to 23th of January) then
      Sample without replacement 300.000 people from total population
      Remove sampled 300.000 people from total population
      Save people who leave Wuhan on day i and are infected
      Look up infection time for each infected traveler in List of infected
      For those sample incubation period
      Include when symptom onset is not before day of travel
      forward time[[i]] <- (time of infection + incubation period) - i
   end if
   New cases of this day <- round(Total cases * Growth rate)
   Total cases <- Total cases + New cases for this day
```

5.2 Simulation study settings

The outbreak in Wuhan was simulated following algorithm 3 and thereby, incubation periods estimates were obtained. The parametric models used for the distribution of incubation period in the simulation settings by Deng were replicated (i.e., gamma distribution ($k = 5, \theta = 0.8$), Weibull distribution ($\lambda = 2, k = 8$), and log-normal distribution ($\mu = 1.8, \sigma^2 = 0.4^2$)). The growth rate was set to 0.14 (Dorigatti et al., 2020), which corresponds to a doubling time of around 5 days. Furthermore, the sample size per outbreak was set to vary between 600, 1200, and 1800, while the additional percentage of infected at day of travel π was set to 0, 0.1, or 0.2. In addition, all the simulation settings were repeated without additional infections at day of travel. This change turns the likelihood from a mixture distribution of forward time and complete incubation period to a distribution of just forward time. For each setting, the outbreak was simulated 1000 times.

5.3 Results

The complete results of the simulation studies can be found in Section 8.3.1. The differences for different sample sizes, growth rate, and values for π are negligible. The same holds true for differences between Deng's and Qin's method. Further, as expected, a larger sample size results in smaller mean squared errors for α, β, and π. See Appendix 8.3.3 for the results when the probability of getting infected at day of travel π is removed from the analysis. The estimates indicate that larger incubation period estimates result. Table 5.1 compares results based on the mixture distribution when $\pi = 0.2$ and the sole forward times (i.e., without π). The three parametric distributions used by Deng were used again. Estimates, true values, and mean squared error for the parameter of the respective distribution and value of π are reported. The 95% confidence intervals are shown in brackets. `R`-code written for the simulation study is available on GitHub: `https://github.com/lnilsnils/SimulationStudy_IncubationTime`. In addition, the code is provided in Appendix 8.4. All analysis were done in `R`-Studio version

4.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

Table 5.1: Novel simulation study results for different distributions

Gamma		Without π		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\alpha}$	5	8.78 (8.71-8.84)	8.75 (8.68-8.82)	4.63 (4.59-4.67)	4.61 (4.57-4.65)
$\hat{\beta}$	0.8	1.11 (1.10-1.12)	1.10 (1.10-1.11)	0.77 (0.77-0.78)	0.77 (0.76-0.77)
$\hat{\pi}$		-	-	0.52 (0.52-0.53)	0.52 (0.51-0.52)
Mean	6.25	7.91 (7.89-7.92)	7.91 (7.89-7.92)	5.55 (5.53-5.57)	5.56 (5.54-5.58)
0.25Q	4.21	5.97 (5.96-5.99)	5.97 (5.96-5.99)	3.94 (3.92-3.96)	3.95 (3.93-3.97)
Median	5.84	7.60 (7.59-7.62)	7.61 (7.60-7.63)	5.56 (5.50-5.63)	5.58 (5.51-5.64)
0.75Q	7.84	9.51 (9.50-9.53)	9.52 (9.51-9.54)	7.56 (7.54-7.57)	7.58 (7.56-7.59)
0.90Q	9.99	11.48 (11.46-11.50)	11.50 (11.48-11.51)	9.73 (9.71-9.74)	9.75 (9.74-9.77)
0.95Q	11.44	12.77 (12.75-12.79)	2.79 (12.77-12.81)	11.2 (11.18-11.21)	11.23 (11.21-11.25)
0.99Q	14.51	15.35 (15.25-15.46)	15.38 (15.28-15.49)	14.32 (14.29-14.34)	14.37 (14.34-14.39)
MSE $\hat{\alpha}$		15.41	15.19	0.57	0.59
MSE $\hat{\beta}$		0.11	0.11	0.01	0.01
MSE $\hat{\pi}$		-	-	0.11	0.11
Weibull		Without π		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\lambda}$	2	2.70 (2.68-2.71)	2.70 (2.68-2.71)	1.98 (1.97-2.00)	2.12 (2.11-2.14)
$\hat{k}$	8	10.31 (10.29-10.33)	10.31 (10.29-10.34)	7.66 (7.61-7.71)	8.15 (8.11-8.19)
$\hat{\pi}$		-	-	0.45 (0.44-0.46)	0.34 (0.33-0.35)
Mean	7.09	9.17 (9.15-9.19)	9.17 (9.15-9.19)	6.80 (6.76-6.84)	6.80 (6.76-6.84)
0.25Q	4.29	6.49 (6.46-6.51)	6.49 (6.46-6.51)	4.09 (4.04-4.13)	4.53 (4.49-4.57)
Median	6.66	8.99 (8.97-9.02)	9.00 (8.97-9.02)	6.36 (6.32-6.41)	6.85 (6.81-6.89)
0.75Q	9.42	11.64 (11.62-11.66)	11.65 (11.63-11.67)	9.04 (8.99-9.09)	9.51 (9.47-9.55)
0.90Q	12.14	14.07 (14.05-14.09)	14.08 (14.06-14.10)	11.70 (11.66-11.74)	12.09 (12.05-12.12)
0.95Q	13.85	15.52 (15.50-15.55)	15.54 (15.51-15.56)	13.38 (13.34-13.42)	13.69 (13.65-13.73)
0.99Q	17.17	18.23 (18.19-18.27)	18.25 (18.21-18.29)	16.67 (16.63-16.71)	16.79 (16.74-16.83)
MSE $\hat{\lambda}$		0.54	0.54	0.06	0.05
MSE $\hat{k}$		5.45	5.45	0.70	0.42
MSE $\hat{\pi}$		-	-	0.23	0.13
Log-normal		Without π		With $\pi = 0.2$	
$m = 1800$	True	Deng's method	Qin's method	Deng's method	Qin's method
$\hat{\mu}$	1.8	2.05 (2.04-2.05)	2.05 (2.04-2.05)	1.72 (1.71-1.72)	1.72 (1.72-1.73)
$\hat{\sigma}$	0.4	0.32 (0.32-0.32)	0.32 (0.32-0.32)	0.43 (0.43-0.43)	0.43 (0.43-0.43)
$\hat{\pi}$		-	-	0.40 (0.39-0.40)	0.39 (0.39-0.40)
Mean	6.25	8.14 (8.13-8.15)	8.14 (8.13-8.15)	6.12 (6.10-6.14)	6.12 (6.10-6.14)
0.25Q	4.21	6.25 (6.23-6.27)	6.25 (6.24-6.27)	4.19 (4.16-4.21)	4.22 (4.19-4.24)
Median	5.84	7.74 (7.72-7.75)	7.74 (7.73-7.76)	5.58 (5.56-5.60)	5.62 (5.59-5.64)
0.75Q	7.84	9.58 (9.57-9.60)	9.59 (9.58-9.60)	7.44 (7.42-7.46)	7.48 (7.46-7.50)
0.90Q	9.99	11.62 (11.60-11.63)	11.63 (11.61-11.65)	9.65 (9.63-9.67)	9.69 (9.67-9.71)
0.95Q	11.44	13.04 (13.01-13.06)	13.06 (13.03-13.08)	11.27 (11.25-11.30)	11.32 (11.30-11.34)
0.99Q	14.51	16.19 (16.15-16.23)	16.22 (16.18-16.26)	15.10 (15.06-15.13)	15.15 (15.11-15.18)
MSE $\hat{\mu}$		0.06	0.06	0.01	0.01
MSE $\hat{\sigma}$		0.01	0.01	0.00	0.00
MSE $\hat{\pi}$		-	-	0.17	0.17

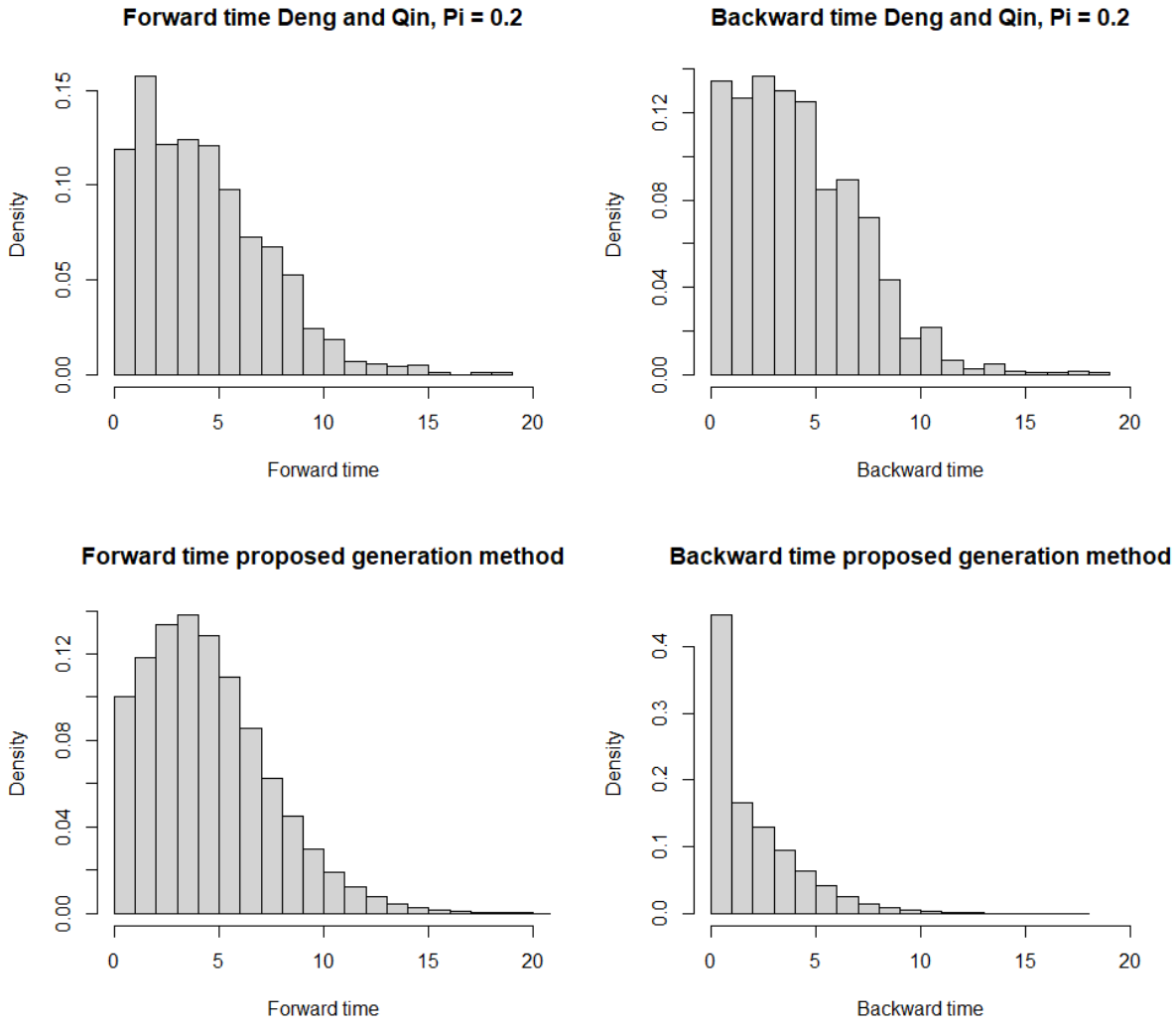


Figure 5.1: Comparison of forward and backward times based on Deng's and Qin's and the exponential data generation methods. Plots are based on 1000 repetitions of Algorithm 1 (Deng and Qin) and Algorithm 3 (proposed data generation method).

Chapter 6

Discussion

Following Qin et al. (2020), Deng et al. (2020) proposed a method to estimate incubation period based on the theory of the renewal process. Both applied their method to data of the initial SARS-CoV-2 outbreak in Wuhan. In contrast to Qin, Deng's method includes interval censoring around the daily reports for symptoms onset results.

The approaches by Deng and Qin offer a number of advantages over existing methods. Contact tracing data that can provides the date of infection tends to be difficult to obtain. The renewal process approaches only require date of symptom onset and travel information. This way, a larger sample sizes can be utilized. This is particularly important at the beginning of an infectious disease outbreak, when timely estimates of the incubation period are critical for containment and analysis of the behaviour of the virus. However, no prior studies have set out to study the validity of the renewal process approach for either of the two methods.

This thesis focus on the distribution of backward time (i.e., time from infection to departure). Deng and Qin assumed a uniformly distributed backward time, while in this thesis a data generation method is proposed that produces backward times that are rather exponentially increasing. It is argued that in this way the sharp increase in the number of infections during an outbreak, as in Wuhan in January 2020, is better captured. With help of a simulation study, this thesis investigated how appropriate Deng's and Qin's approaches are for this type of data.

The obtained estimates reflect previous published estimates in the literature (e.g., Linton et al. (2020), Lauer et al. (2020), Backer et al. (2020)). However, the results do not match the estimates of Deng and Qin. The difference in data generation, leads to an estimated incubation time that is about 1.5-2 days shorter. Deng claims that the longer estimates obtained with their methods can be explained by the longer follow-up periods after leaving Wuhan. This way, they don't fall pray to the selection bias, as longer incubation periods are just as likely to be selected as shorter ones. Like Deng and Qin, this thesis considers a follow-up period of 25 days after departure. This suggests that the differences between the results cannot be explained by the follow-up time, but rather by the differences in data generation and the change from an interval-censored to a renewal process-based method.

Another important finding is that the estimated mean and median incubation period is around 7-9 days (depending on the distribution) when π is removed. The estimates are based on the proposed data generation method. The estimates are, therefore, about 2-2.5 days shorter when π is included. Furthermore, as mentioned above, the estimates are 1.5-2 days shorter when π is included than when a uniformly distributed backward time is assumed. Thus, the results of this thesis suggest that Qin and Deng introduced with π and the uniformly distributed backward time two type of biases, which partly channel each other out.

Lastly, contrary to expectations, the simulation study returned an estimated value for π between 0.40 and 0.50. This is even the case when π is set to 0; no additional infections at day of travel are added. In this scenario, all forward times are drawn from the exponential data generation method. This may be explained by the fact that most infected people were

not included in the sample because they developed symptoms before departure, whereas people who were infected at day of departure were automatically included in the sample. The large percentage of people who were infected at day of travel, can also be seen in the bottom-right plot in Figure 5.1.

6.1 Limitations

A note of caution is required due to possible violation of assumptions. Firstly, Figure 5.1 depicts the forward and backward times based on Deng and Qin and the exponential data generation method. Following the theory of the renewal process, one would expect the distributions of the forward and backward time to be a mirrored versions of each other. Furthermore, according to the assumptions of the renewal process, the densities of the forward times are expected to be decreasing functions, proportional to the survival function of incubation period (De Gruttola & Lagakos, 1989). Since the generates data in Algorithm 3 as well as Deng's and Qin's method violate these assumptions, the estimates obtained via the renewal process method should be interpreted with caution.

The fact that the estimates obtained via the exponential data generation method are consistent with most of the literature is unexpected: firstly, given the violated assumptions, secondly, due to the inclusion of additional infections on travel day (i.e., π). The latter because it is not clear whether the correction for additional infections at travel day are is warranted. Deng justifies the inclusion of π by saying that crowded airports increase the risk of infection on the day of departure. One could argue that a crowded mega-city like Wuhan, at a time when there were no restrictions, offered no less opportunities to become infected.

In addition, it is important to consider that that the assumption that individuals were infected either in Wuhan or on the day of departure from Wuhan is likely to be violated if real data are used. As mentioned by Deng, people might also get infected by other members in their tour group or family after departure.It should be noted that this is not the case in the simulation study used in this work, as this scenario was not investigated.

Another source of uncertainty are the values that were assumed for the simulated outbreak in Wuhan (see section 5.1). We assumed 125 SARS-CoV-2 at 5th January 2020. This date marks the start of the simulated outbreak. Official reports from the Wuhan Municipal Health Commission state 59 known cases in the Wuhan region at January 5. However, official numbers of SARS-CoV-2 infections at the begin of the pandemic are assumed to suffer from under-reporting.

Furthermore, no reliable data could be obtained that could provide information on the number of people who entered or left Wuhan the days prior to the lockdown. To arrive at about the same sample size ($\sim$ 1200) as Deng per simulated outbreak, 245.000 are believed to have left/entered the city daily prior to January 19. In addition, it was assumed that daily 490.000 people left between the January 19 and January 23. Appendix 8.3.2 shows the results when additional simulation studies are performed when alternative numbers assumed. The differences with the original assumed numbers are negligible.

6.2 Future Research

The major advantage of the renewal process approach over existing methods is its ability to provide estimates based on data which are rather easy to collect. In addition, Deng's and Qin's method highlight the need of having a method that can provide robust incubation period estimates at the beginning of an outbreak, such as in Wuhan. Similarly, an adapted version that excludes π and assumes the uniform distributed backward time can be helpful in an ongoing epidemic.

As outline in this discussion, the assumptions underlying the renewal process are violated by Deng's and Qin's data and by the data based on the novel data generation method (Algorithm

3). However, is it still unclear whether the theory of renewal process could be correctly used to estimate incubation periods. Further studies are needed to understand why the assumptions of the renewal process are violated in the studies in which the methods have been used so far.

6.3 Conclusion

Deng and Qin proposed to employ the renewal process approach to estimate incubation periods. Deng and Qin assumed that the time between infection and departure is uniformly distributed, even so the number of SARS-CoV-2 infection spread exponentially. The proposed data generation method took the exponential increase of cases into account. The obtained estimates from the simulation studies are in line with the majority of estimates in the literature. However, the assumptions underlying the renewal process are violated in this thesis as well as in Deng's and Qin's study. Hence, Deng's, Qin's and the method proposed as part of this thesis should be used with caution.

Applying the theory of the renewal process to estimating incubation periods appears, nonetheless, to be a promising idea. To get a complete picture, further studies are needed to examine the reasons for the violated assumptions in more detail. The renewal process framework may be an attractive alternative to commonly used interval-censored methods. They may help to use larger sample sizes and provide estimates even so no information on the date of infection is available. This is turn, may help to utilize larger sample sizes for incubation period estimates in future infectious disease outbreaks.

Chapter 7

References

Backer, J. A., Klinkenberg, D., & Wallinga, J. (2020). Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, china, 20–28 january 2020. *Eurosurveillance*, *25*(5), 2000062.

Cox, D. (1964). Dr cox, renewal theory. *The Annals of Mathematical Statistics*, *35*(1), 448–449.

De Gruttola, V., & Lagakos, S. W. (1989). Analysis of doubly-censored survival data, with application to aids. *Biometrics*, 1–11.

Deng, Y., You, C., Liu, Y., Qin, J., & Zhou, X.-H. (2020). Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for covid-19 outbreak in china. *Biometrics*.

Dietz, K., & Heesterbeek, J. (2000). Bernoulli was ahead of modern epidemiology. *Nature*, *408*(6812), 513–514.

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, *20*(5), 533–534.

Dorigatti, I., Okell, L., Cori, A., Imai, N., Baguelin, M., Bhatia, S., ... others (2020). Report 4: severity of 2019-novel coronavirus (ncov). *Imperial College London, London*.

Ghebreyesus, T. (2020). *Who director-general's opening remarks at the media briefing on covid-19-2020.*

Groeneboom, P. (2021). Estimation of the incubation time distribution for covid-19. *Statistica Neerlandica*, *75*(2), 161–179.

Jung, S.-m., Akhmetzhanov, A. R., Hayashi, K., Linton, N. M., Yang, Y., Yuan, B., ... Nishiura, H. (2020). Real-time estimation of the risk of death from novel coronavirus (covid-19) infection: inference using exported cases. *Journal of clinical medicine*, *9*(2), 523.

Krämer, A., Kretzschmar, M., & Krickeberg, K. (2010). *Modern infectious disease epidemiology: Concepts, methods, mathematical models, and public health.* Springer.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., ... Lessler, J. (2020). The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, *172*(9), 577–582.

Linton, N. M., Kobayashi, T., Yang, Y., Hayashi, K., Akhmetzhanov, A. R., Jung, S.-m., ... Nishiura, H. (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *Journal of clinical medicine*, *9*(2), 538.

Nishiura, H. (2007). Early efforts in modeling the incubation period of infectious diseases with an acute course of illness. *Emerging themes in epidemiology*, *4*(1), 1–12.

Qin, J., You, C., Lin, Q., Hu, T., Yu, S., & Zhou, X.-H. (2020). Estimation of incubation period distribution of covid-19 using disease onset forward time: a novel cross-sectional and forward follow-up study. *Science advances*, *6*(33), eabc1202.

Reich, N. G., Lessler, J., Cummings, D. A., & Brookmeyer, R. (2009). Estimating incubation period distributions with coarse data. *Statistics in medicine*, *28*(22), 2769–2784.

Voysey, M., Clemens, S. A. C., Madhi, S. A., Weckx, L. Y., Folegatti, P. M., Aley, P. K., ... others (2021). Safety and efficacy of the chadox1 ncov-19 vaccine (azd1222) against sars-cov-2: an interim analysis of four randomised controlled trials in brazil, south africa, and the uk. *The Lancet*, *397*(10269), 99–111.

Wassie, G. T., Azene, A. G., Bantie, G. M., Dessie, G., & Aragaw, A. M. (2020). Incubation period of severe acute respiratory syndrome novel coronavirus 2 that causes coronavirus disease 2019: A systematic review and meta-analysis. *Current Therapeutic Research, Clinical and Experimental*, *93*.

World Bank, T. (2020). The global economic outlook during the covid-19 pandemic: A changed world. *World Bank Feature Story*.

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... others (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-ncov) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International journal of infectious diseases*, *92*, 214–217.

Chapter 8

Appendix

8.1 Additional computations Section 2.2

8.1.1 Details about computations for the forward time density function (2.12)

The density of the forward time V is as follows

$$g(v) = \frac{\bar{F}(v)}{\mu}, \quad v \geq 0$$

where μ is the mean incubation period and $\bar{F}(v)$ is the survival function. Qin et al. (2020) assume that the forward time V follows a Weibull distribution with parameters α and λ. Given that the survival function for Weibull distribution is $\exp\{-(\lambda v)^\alpha\}$ and the mean for the Weibull distribution $\frac{\Gamma(1+1/\alpha)}{\lambda}$, the density of the forward time can be expressed as:

$$\begin{aligned} g(v) &= \frac{\exp\{-(\lambda v)^\alpha\}}{\frac{\Gamma(1+\frac{1}{\alpha})}{\lambda}} \\ &= \lambda \frac{\exp\{-(\lambda v)^\alpha\}}{\Gamma(1+\frac{1}{\alpha})} \\ &= \lambda \frac{\exp\{-(\lambda v)^\alpha\}}{\frac{1}{\alpha}\Gamma(\frac{1}{\alpha})} \\ &= \alpha\lambda \frac{\exp\{-(\lambda v)^\alpha\}}{\Gamma(\frac{1}{\alpha})} \end{aligned}$$

8.1.2 Details about computations for the likelihood (2.16)

$$f(y) = \alpha\lambda(y\lambda)^{\alpha-1}\exp\{-(y\lambda)^\alpha\}, \quad y \geq 0$$

$$\begin{aligned} h(v) &= (1-\pi)g(v) + \pi f(v), \qquad v \geq 0 \\ &= (1-\pi)\alpha\lambda \frac{\exp\{-(\lambda\nu)^\alpha\}}{\Gamma(\frac{1}{\alpha})} + \pi\alpha\lambda(\nu\lambda)^{\alpha-1}\exp\{-(\nu\lambda)^\alpha\} \\ &= \alpha\lambda(\pi(\nu\lambda)^{\alpha-1}\exp\{-(\lambda\nu)^\alpha\} + \frac{(1-\pi)\exp\{-(\lambda\nu)^\alpha\}}{\Gamma(\frac{1}{\alpha})} \\ &= \alpha\lambda(\pi(\nu\lambda)^{\alpha-1}\exp\{-(\lambda\nu)^\alpha\} + \frac{(1-\pi)}{\Gamma(\frac{1}{\alpha})}\exp\{-(\lambda\nu)^\alpha\} \\ &= \alpha\lambda\{\pi(v\lambda)^{\alpha-1} + \frac{(1-\pi)}{\Gamma(\frac{1}{\alpha})}\}\exp(-(v\lambda)^\alpha) \end{aligned}$$

8.2 Supplementary material Deng et al. (2020)

8.2.1 Likelihood for mixture distribution

The latent variable δ is introduced to derive the likelihood of the mixture distribution, which denotes whether the jth person got infected at departure ($\delta = 1$) or before ($\delta = 0$). The density of a mixture distribution when accounting for the probability π of getting infected at day of travel is given as follows:

$$\delta_j \sim Bin(1, \pi), j = 1, .., m,$$
$$t_j|(\delta_j = 1) \sim f_I^p(\cdot; \theta),\ t_j|(\delta_j = 0) \sim h^p(\cdot; \theta),$$

The conditional likelihood is

$$L(\theta; t_1, ..., t_m|\delta_1, ..., \delta_m) = \prod_{j=1}^{m} \{f_I^p(t_j; \theta)\}^{\delta_j} \{h^p(t_j; \theta)\}^{1-\delta_j}$$

The marginal likelihood can be obtained by integrating out the latent variable $\{\delta_1 ... \delta_m\}$

$$L(\theta, \pi; t_1, ..., t_m) = \prod_{j=1}^{m} \{\pi f_I^p(t_j; \theta) + (1 - \pi) h^p(t_j; \theta)\}$$

where f_1 represents the PDF of the incubation period and h the PDF of forward time.

8.3 Results simulation study

8.3.1 Incubation period estimates

Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.77	(4.68 -4.87)	4.78	(4.69-4.88)
$\hat{\beta}$	0.80	(0.78-0.81)	0.79	(0.78-0.80)
$\hat{\pi}$	0.41	(0.40-0.42)	0.40	(0.40-0.41)
Mean	5.90	(5.86-5.94)	5.90	(5.86-5.94)
0.25Q	3.89	(3.85-3.93)	3.91	(3.87-3.95)
Median	5.47	(5.43-5.51)	5.5	(5.46-5.54)
0.75Q	7.45	(7.41-7.49)	7.48	(7.44-7.51)
0.90Q	9.59	(9.55-9.62)	9.62	(9.59-9.66)
0.95Q	11.04	(11.01-11.08)	11.08	(11.04-11.11)
0.99Q	14.13	(14.09-14.18)	14.17	(14.13-14.21)
MSE $\hat{\alpha}$	2.34	-	2.29	-
MSE $\hat{\beta}$	0.03	-	0.03	-
MSE $\hat{\pi}$	0.19	-	0.18	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.62	(4.55-4.70)	4.63	(4.55-4.70)
$\hat{\beta}$	0.78	(0.77-0.79)	0.78	(0.77-0.79)
$\hat{\pi}$	0.42	(0.41-0.43)	0.41	(0.41-0.42)
Mean	5.85	(5.82-5.89)	5.85	(5.82-5.89)
0.25Q	3.84	(3.80-3.87)	3.85	(3.82-3.89)
Median	5.42	(5.39-5.46)	5.44	(5.41-5.48)
0.75Q	7.4	(7.37-7.43)	7.43	(7.40-7.46)
0.90Q	9.55	(9.52-9.58)	9.58	(9.55-9.61)
0.95Q	11.01	(10.98-11.04)	11.04	(11.01-11.07)
0.99Q	14.11	(14.07-14.14)	14.15	(14.12-14.18)
MSE $\hat{\alpha}$	1.64	-	1.62	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.19	-	0.18	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.53	(4.47-4.59)	4.53	(4.47-4.59)
$\hat{\beta}$	0.77	(0.76-0.78)	0.77	(0.76 -0.77)
$\hat{\pi}$	0.42	(0.41-0.42)	0.41	(0.41-0.42)
Mean	5.85	(5.82-5.87)	5.85	(5.82-5.87)
0.25Q	3.82	(3.79-3.85)	3.83	(3.80-3.86)
Median	5.41	(5.38-5.44)	5.43	(5.40-5.46)
0.75Q	7.40	(7.37-7.43)	7.42	(7.40-7.45)
0.90Q	9.56	(9.53-9.58)	9.59	(9.56-9.61)
0.95Q	11.03	(11.00-11.05)	11.06	(11.04-11.08)
0.99Q	14.15	(14.12-14.18)	14.19	(14.16-14.22)
MSE $\hat{\alpha}$	1.23	-	1.23	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.19	-	0.18	-

Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.77	(4.69-4.85)	4.77	(4.70-4.85)
$\hat{\beta}$	0.79	(0.78-0.80)	0.79	(0.78-0.80)
$\hat{\pi}$	0.46	(0.46-0.47)	0.46	(0.45-0.47)
Mean	5.96	(5.92-5.99)	5.96	(5.92-5.99)
0.25Q	3.94	(3.90-3.97)	3.95	(3.92-3.99)
Median	5.53	(5.49-5.56)	5.55	(5.52-5.58)
0.75Q	7.51	(7.48-7.54)	7.54	(7.51-7.57)
0.90Q	9.66	(9.63-9.69)	9.69	(9.66-9.72)
0.95Q	11.12	(11.08-11.15)	11.15	(11.12-11.18)
0.99Q	14.21	(14.17-14.25)	14.25	(14.21-14.29)
MSE $\hat{\alpha}$	1.64	-	1.63	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.15	-	0.14	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.55	(4.50-4.60)	4.54	(4.49-4.59)
$\hat{\beta}$	0.77	(0.76 - 0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.47	(0.46-0.48)	0.47	(0.46-0.47)
Mean	5.91	(5.89-5.93)	5.91	(5.89-5.93)
0.25Q	3.87	(3.84-3.89)	3.88	(3.85-3.9)
Median	5.47	(5.45-5.5)	5.49	(5.46-5.51)
0.75Q	7.47	(7.45-7.50)	7.49	(7.47-7.52)
0.90Q	9.64	(9.62-9.66)	9.67	(9.65-9.69)
0.95Q	11.11	(11.09-11.14)	11.15	(11.13-11.17)
0.99Q	14.24	(14.22-14.27)	14.29	(14.26-14.32)
MSE $\hat{\alpha}$	0.87	-	0.88	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.15	-	0.14	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.56	(4.52-4.61)	4.55	(4.50- 4.60)
$\hat{\beta}$	0.77	(0.76-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.47	(0.46-0.47)	0.46	(0.46-0.47)
Mean	5.91	(5.89-5.93)	5.91	(5.89-5.93)
0.25Q	3.88	(3.85-3.9)	3.88	(3.86-3.91)
Median	5.48	(5.46-5.5)	5.49	(5.47-5.51)
0.75Q	7.48	(7.46-7.5)	7.5	(7.48-7.52)
0.90Q	9.64	(9.62-9.66)	9.67	(9.65-9.69)
0.95Q	11.11	(11.09-11.13)	11.14	(11.13-11.16)
0.99Q	14.23	(14.21-14.26)	14.28	(14.25-14.31)
MSE $\hat{\alpha}$	0.76	-	0.77	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.14	-	0.14	-

Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.78	(4.72-4.84)	4.77	(4.71-4.84)
$\hat{\beta}$	0.79	(0.78-0.80)	0.79	(0.78-0.79)
$\hat{\pi}$	0.52	(0.51-0.52)	0.51	(0.50-0.52)
Mean	6.01	(5.99-6.04)	6.01	(5.99-6.04)
0.25Q	3.98	(3.96-4.01)	3.99	(3.97-4.02)
Median	5.59	(5.56-5.61)	5.6	(5.58-5.63)
0.75Q	7.58	(7.55-7.6)	7.6	(7.58-7.63)
0.90Q	9.73	(9.71-9.75)	9.76	(9.74-9.78)
0.95Q	11.19	(11.17-11.21)	11.22	(11.2-11.25)
0.99Q	14.28	(14.25-14.32)	14.33	(14.3-14.36)
MSE $\hat{\alpha}$	1.08	-	1.07	-
MSE $\hat{\beta}$	0.02	-	0.01	-
MSE $\hat{\pi}$	0.11	-	0.11	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.71	(4.66-4.75)	4.69	(4.65-4.74)
$\hat{\beta}$	0.78	(0.78-0.79)	0.78	(0.77-0.78)
$\hat{\pi}$	0.52	(0.51-0.52)	0.51	(0.51-0.52)
Mean	6.00	(5.98-6.02)	6.00	(5.98-6.02)
0.25Q	3.97	(3.94-3.99)	3.97	(3.95-3.99)
Median	5.57	(5.55-5.59)	5.58	(5.56-5.60)
0.75Q	7.57	(7.55-7.59)	7.59	(7.57-7.61)
0.90Q	9.72	(9.70-9.74)	9.75	(9.73-9.77)
0.95Q	11.18	(11.16-11.2)	11.22	(11.2-11.24)
0.99Q	14.28	(14.25-14.31)	14.33	(14.30-14.36)
MSE $\hat{\alpha}$	0.67	-	0.69	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.11	-	0.11	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.63	(4.59-4.67)	4.61	(4.57-4.65)
$\hat{\beta}$	0.77	(0.77-0.78)	0.77	(0.76-0.77)
$\hat{\pi}$	0.52	(0.52-0.53)	0.52	(0.51-0.52)
Mean	5.98	(5.97-6.00)	5.98	(5.97-6.00)
0.25Q	3.94	(3.92-3.96)	3.95	(3.93-3.97)
Median	5.55	(5.53-5.57)	5.56	(5.54-5.58)
0.75Q	7.56	(7.54-7.57)	7.58	(7.56-7.59)
0.90Q	9.73	(9.71-9.74)	9.75	(9.74-9.77)
0.95Q	11.2	(11.18-11.21)	11.23	(11.21-11.25)
0.99Q	14.32	(14.29-14.34)	14.37	(14.34-14.39)
MSE $\hat{\alpha}$	0.57	-	0.59	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.11	-	0.11	-

Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	2.03	(2.00-2.05)	2.16	(2.15-2.18)
$\hat{k}$	7.79	(7.72-7.85)	8.28	(8.22-8.33)
$\hat{\pi}$	0.43	(0.42-0.45)	0.32	(0.31-0.33)
Mean	6.92	(6.86-6.98)	6.92	(6.86-6.98)
0.25Q	4.21	(4.15-4.28)	4.65	(4.60-4.70)
Median	6.5	(6.43-6.56)	6.98	(6.93-7.04)
0.75Q	9.16	(9.10-9.23)	9.63	(9.58-9.69)
0.90Q	11.81	(11.75-11.87)	12.2	(12.15-12.26)
0.95Q	13.48	(13.42-13.53)	13.8	(13.74-13.85)
0.99Q	16.74	(16.69-16.79)	16.87	(16.81-16.93)
MSE $\hat{\lambda}$	0.11	-	0.11	-
MSE $\hat{k}$	1.15	-	0.85	-
MSE $\hat{\pi}$	0.23	-	0.13	-
Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	1.99	(1.97-2.01)	2.14	(2.12-2.15)
$\hat{k}$	7.68	(7.63-7.73)	8.17	(8.13-8.22)
$\hat{\pi}$	0.44	(0.43-0.46)	0.33	(0.33-0.34)
Mean	6.82	(6.78-6.87)	6.82	(6.78-6.87)
0.25Q	4.11	(4.06-4.16)	4.56	(4.52-4.6)
Median	6.39	(6.34-6.44)	6.88	(6.84 -6.93)
0.75Q	9.06	(9.01-9.11)	9.53	(9.49-9.57)
0.90Q	11.71	(11.67-11.76)	12.1	(12.06-12.14)
0.95Q	13.39	(13.34-13.43)	13.7	(13.66-13.74)
0.99Q	16.66	(16.62-16.70)	16.78	(16.74-16.83)
MSE $\hat{\lambda}$	0.07	-	0.07	-
MSE $\hat{k}$	0.80	-	0.50	-
MSE $\hat{\pi}$	0.23	-	0.13	-
Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	1.98	(1.97-2.00)	2.12	(2.11-2.14)
$\hat{k}$	7.66	(7.61-7.71)	8.15	(8.11-8.19)
$\hat{\pi}$	0.45	(0.44-0.46)	0.34	(0.33-0.35)
Mean	6.80	(6.76-6.84)	6.80	(6.76-6.84)
0.25Q	4.09	(4.04-4.13)	4.53	(4.49-4.57)
Median	6.36	(6.32-6.41)	6.85	(6.81-6.89)
0.75Q	9.04	(8.99-9.09)	9.51	(9.47-9.55)
0.90Q	11.70	(11.66-11.74)	12.09	(12.05-12.12)
0.95Q	13.38	(13.34-13.42)	13.69	(13.65-13.73)
0.99Q	16.67	(16.63-16.71)	16.79	(16.74-16.83)
MSE $\hat{\lambda}$	0.06	-	0.05	-
MSE $\hat{k}$	0.70	-	0.42	-
MSE $\hat{\pi}$	0.23	-	0.13	-

Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	2.02	(2.00-2.03)	2.15	(2.13-2.16)
$\hat{k}$	7.80	(7.74-7.86)	8.24	(8.19-8.29)
$\hat{\pi}$	0.49	(0.48-0.50)	0.39	(0.38-0.40)
Mean	6.93	(6.88-6.98)	6.93	(6.88-6.98)
0.25Q	4.21	(4.15-4.26)	4.61	(4.57-4.65)
Median	6.50	(6.44-6.56)	6.94	(6.89-6.99)
0.75Q	9.18	(9.13-9.24)	9.60	(9.55-9.65)
0.90Q	11.84	(11.79-11.89)	12.18	(12.13-12.23)
0.95Q	13.52	(13.47-13.56)	13.78	(13.73-13.83)
0.99Q	16.79	(16.74-16.84)	16.87	(16.82-16.92)
MSE $\hat{\lambda}$	0.09	-	0.08	-
MSE $\hat{k}$	0.88	-	0.66	-
MSE $\hat{\pi}$	0.19	-	0.10	-
Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	2.00	(1.98-2.01)	2.13	(2.12-2.14)
$\hat{k}$	7.75	(7.70-7.79)	8.18	(8.14-8.22)
$\hat{\pi}$	0.49	(0.49-0.50)	0.40	(0.39-0.40)
Mean	6.88	(6.84-6.92)	6.88	(6.84-6.92)
0.25Q	4.15	(4.11-4.20)	4.56	(4.52-4.59)
Median	6.45	(6.40-6.49)	6.89	(6.85-6.92)
0.75Q	9.13	(9.09-9.18)	9.54	(9.51-9.58)
0.90Q	11.8	(11.76-11.84)	12.13	(12.09-12.16)
0.95Q	13.48	(13.44-13.51)	13.73	(13.70-13.77)
0.99Q	16.76	(16.72-16.8)	16.83	(16.79-16.87)
MSE $\hat{\lambda}$	0.06	-	0.06	-
MSE $\hat{k}$	0.61	-	0.42	-
MSE $\hat{\pi}$	0.18	-	0.10	-
Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	1.97	(1.96-1.98)	2.11	(2.10-2.12)
$\hat{k}$	7.70	(7.66-7.74)	8.14	(8.11-8.18)
$\hat{\pi}$	0.50	(0.49-0.51)	0.40	(0.39-0.40)
Mean	6.83	(6.80-6.87)	6.83	(6.80-6.87)
0.25Q	4.09	(4.06-4.13)	4.51	(4.48-4.54)
Median	6.39	(6.36-6.43)	6.84	(6.81-6.87)
0.75Q	9.09	(9.06-9.13)	9.51	(9.48-9.54)
0.90Q	11.78	(11.75-11.81)	12.11	(12.08-12.14)
0.95Q	13.47	(13.44-13.50)	13.73	(13.70-13.76)
0.99Q	16.78	(16.75-16.82)	16.86	(16.82-16.89)
MSE $\hat{\lambda}$	0.04	-	0.04	-
MSE $\hat{k}$	0.46	-	0.28	-
MSE $\hat{\pi}$	0.18	-	0.10	-

Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	2.01	(1.99-2.03)	2.13	(2.12-2.15)
$\hat{k}$	7.80	(7.76-7.85)	8.20	(8.16-8.24)
$\hat{\pi}$	0.55	(0.54-0.56)	0.45	(0.44-0.46)
Mean	6.93	(6.89-6.97)	6.93	(6.89-6.97)
0.25Q	4.19	(4.15-4.24)	4.57	(4.53-4.61)
Median	6.50	(6.45-6.55)	6.90	(6.86-6.95)
0.75Q	9.19	(9.15-9.24)	9.57	(9.53-9.61)
0.90Q	11.87	(11.82-11.91)	12.16	(12.12-12.20)
0.95Q	13.55	(13.51-13.59)	13.77	(13.73-13.81)
0.99Q	16.84	(16.80-16.88)	16.88	(16.84-16.93)
MSE $\hat{\lambda}$	0.08	-	0.07	-
MSE $\hat{k}$	0.67	-	0.50	-
MSE $\hat{\pi}$	0.15	-	0.08	-
Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	1.99	(1.98-2.00)	2.11	(2.10-2.12)
$\hat{k}$	7.76	(7.73-7.80)	8.15	(8.12-8.18)
$\hat{\pi}$	0.55	(0.54-0.56)	0.46	(0.45-0.46)
Mean	6.89	(6.86-6.92)	6.89	(6.86-6.92)
0.25Q	4.15	(4.12-4.19)	4.52	(4.49-4.54)
Median	6.46	(6.42-6.49)	6.85	(6.82-6.88)
0.75Q	9.15	(9.12-9.19)	9.51	(9.48-9.54)
0.90Q	11.83	(11.8-11.86)	12.11	(12.08-12.14)
0.95Q	13.51	(13.48-13.54)	13.72	(13.69-13.75)
0.99Q	16.8	(16.76-16.83)	16.83	(16.8-16.87)
MSE $\hat{\lambda}$	0.04	-	0.04	-
MSE $\hat{k}$	0.39	-	0.25	-
MSE $\hat{\pi}$	0.08	-	0.08	-
Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\lambda}$	1.98	(1.97-1.99)	2.10	(2.09-2.11)
$\hat{k}$	7.76	(7.73-7.79)	8.15	(8.12-8.17)
$\hat{\pi}$	0.55	(0.54-0.55)	0.46	(0.45-0.46)
Mean	6.88	(6.86-6.91)	6.88	(6.86-6.91)
0.25Q	4.14	(4.11-4.16)	4.50	(4.48-4.53)
Median	6.45	(6.42-6.48)	6.84	(6.82-6.87)
0.75Q	9.16	(9.13-9.19)	9.52	(9.49-9.54)
0.90Q	11.84	(11.82-11.87)	12.12	(12.10-12.15)
0.95Q	13.53	(13.51-13.56)	13.74	(13.72-13.77)
0.99Q	16.84	(16.81-16.87)	16.87	(16.85-16.90)
MSE $\hat{\lambda}$	0.03	-	0.03	-
MSE $\hat{k}$	0.30	-	0.18	-
MSE $\hat{\pi}$	0.13	-	0.07	-

Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.72	(1.72-1.73)	1.73	(1.72-1.74)
$\hat{\sigma}$	0.43	(0.42-0.43)	0.42	(0.42-0.43)
$\hat{\pi}$	0.40	(0.39-0.40)	0.39	(0.38-0.39)
Mean	6.16	(6.13-6.19)	6.16	(6.13-6.19)
0.25Q	4.24	(4.20-4.27)	4.27	(4.23-4.3)
Median	5.63	(5.59-5.66)	5.66	(5.63-5.70)
0.75Q	7.48	(7.45-7.52)	7.52	(7.49-7.56)
0.90Q	9.68	(9.65-9.71)	9.72	(9.69-9.76)
0.95Q	11.29	(11.26-11.33)	11.34	(11.31-11.38)
0.99Q	15.1	(15.05-15.15)	15.15	(15.10-15.20)
MSE $\hat{\mu}$	0.02	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.17	-	0.16	-
Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.72	(1.72-1.73)	1.73	(1.73-1.73)
$\hat{\sigma}$	0.43	(0.42-0.43)	0.42	(0.42-0.43)
$\hat{\pi}$	0.39	(0.39-0.40)	0.38	(0.38-0.39)
Mean	6.15	(6.13-6.17)	6.15	(6.13-6.17)
0.25Q	4.22	(4.20-4.25)	4.25	(4.23-4.28)
Median	5.62	(5.59-5.64)	5.65	(5.63-5.68)
0.75Q	7.48	(7.45-7.50)	7.52	(7.49-7.54)
0.90Q	9.68	(9.65-9.70)	9.72	(9.69-9.74)
0.95Q	11.29	(11.26-11.32)	11.34	(11.31-11.36)
0.99Q	15.09	(15.05-15.13)	15.14	(15.10-15.18)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.16	-	0.16	-
Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.72	(1.71-1.72)	1.72	(1.72-1.73)
$\hat{\sigma}$	0.43	(0.43-0.43)	0.43	(0.43-0.43)
$\hat{\pi}$	0.40	(0.39-0.40)	0.39	(0.39-0.40)
Mean	6.12	(6.10-6.14)	6.12	(6.10-6.14)
0.25Q	4.19	(4.16-4.21)	4.22	(4.19-4.24)
Median	5.58	(5.56-5.60)	5.62	(5.59-5.64)
0.75Q	7.44	(7.42-7.46)	7.48	(7.46-7.50)
0.90Q	9.65	(9.63-9.67)	9.69	(9.67-9.71)
0.95Q	11.27	(11.25-11.30)	11.32	(11.30-11.34)
0.99Q	15.10	(15.06-15.13)	15.15	(15.11-15.18)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.17	-	0.16	-

Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.74	(1.73-1.74)	1.74	(1.74-1.75)
$\hat{\sigma}$	0.42	(0.42-0.42)	0.42	(0.42-0.42)
$\hat{\pi}$	0.45	(0.44-0.45)	0.44	(0.43-0.45)
Mean	6.23	(6.20-6.25)	6.23	(6.20-6.25)
0.25Q	4.30	(4.27-4.33)	4.32	(4.30-4.35)
Median	5.70	(5.68-5.73)	5.73	(5.70-5.75)
0.75Q	7.56	(7.54-7.59)	7.60	(7.57-7.62)
0.90Q	9.76	(9.73-9.78)	9.80	(9.77-9.82)
0.95Q	11.37	(11.34- 11.40)	11.41	(11.38-11.44)
0.99Q	15.15	(15.11-15.20)	15.20	(15.16-15.25)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.13	-	0.12	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.73	(1.73-1.74)	1.74	(1.74-1.74)
$\hat{\sigma}$	0.42	(0.42-0.43)	0.42	(0.42-0.43)
$\hat{\pi}$	0.45	(0.45-0.46)	0.44	(0.44-0.45)
Mean	6.20	(6.19-6.22)	6.20	(6.19-6.22)
0.25Q	4.26	(4.25-4.28)	4.29	(4.27-4.30)
Median	5.67	(5.65-5.69)	5.70	(5.68-5.72)
0.75Q	7.54	(7.52-7.56)	7.58	(7.56-7.59)
0.90Q	9.75	(9.74-9.77)	9.79	(9.77-9.81)
0.95Q	11.38	(11.36-11.40)	11.42	(11.40-11.44)
0.99Q	15.2	(15.16-15.23)	15.25	(15.21-15.28)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.13	-	0.12	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.73	(1.73-1.74)	1.74	(1.74-1.74)
$\hat{\sigma}$	0.42	(0.42-0.42)	0.42	(0.42-0.42)
$\hat{\pi}$	0.45	(0.45-0.45)	0.44	(0.44-0.45)
Mean	6.20	(6.19-6.22)	6.20	(6.19-6.22)
0.25Q	4.27	(4.25-4.28)	4.29	(4.27-4.30)
Median	5.67	(5.65-5.69)	5.7	(5.68-5.71)
0.75Q	7.54	(7.52-7.56)	7.57	(7.55-7.59)
0.90Q	9.75	(9.73-9.76)	9.78	(9.77-9.80)
0.95Q	11.36	(11.34-11.38)	11.41	(11.39-11.43)
0.99Q	15.17	(15.13-15.20)	15.22	(15.18-15.25)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.13	-	0.12	-

Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.75	(1.75-1.76)	1.76	(1.75-1.76)
$\hat{\sigma}$	0.42	(0.41-0.42)	0.42	(0.41-0.42)
$\hat{\pi}$	0.50	(0.50-0.51)	0.50	(0.49-0.50)
Mean	6.30	(6.28-6.32)	6.30	(6.28-6.32)
0.25Q	4.37	(4.34-4.39)	4.39	(4.36-4.41)
Median	5.77	(5.75 - 5.8)	5.80	(5.78-5.82)
0.75Q	7.64	(7.62-7.66)	7.67	(7.65-7.69)
0.90Q	9.84	(9.82-9.86)	9.88	(9.85-9.90)
0.95Q	11.45	(11.42-11.47)	11.49	(11.46-11.51)
0.99Q	15.21	(15.17-15.25)	15.26	(15.22-15.31)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.10	-	0.09	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.75	(1.75-1.75)	1.75	(1.75-1.76)
$\hat{\sigma}$	0.42	(0.42-0.42)	0.42	(0.42-0.42)
$\hat{\pi}$	0.50	(0.50-0.51)	0.50	(0.49-0.50)
Mean	6.28	(6.27-6.29)	6.28	(6.27-6.29)
0.25Q	4.34	(4.32-4.35)	4.36	(4.34-4.37)
Median	5.75	(5.74-5.77)	5.77	(5.76-5.79)
0.75Q	7.63	(7.61-7.64)	7.66	(7.64-7.67)
0.90Q	9.84	(9.82-9.86)	9.88	(9.86-9.89)
0.95Q	11.46	(11.44-11.48)	11.5	(11.48-11.52)
0.99Q	15.25	(15.22-15.29)	15.31	(15.27-15.34)
MSE $\hat{\mu}$	0.00	-	0.00	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.10	-	0.09	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(2.5%CI-0.975%CI)	Estimate	(2.5%CI-0.975%CI)
$\hat{\mu}$	1.75	(1.75-1.75)	1.75	(1.75-1.75)
$\hat{\sigma}$	0.42	(0.42-0.42)	0.42	(0.42-0.42)
$\hat{\pi}$	0.50	(0.50-0.51)	0.50	(0.49-0.50)
Mean	6.28	(6.26-6.29)	6.28	(6.26-6.29)
0.25Q	4.34	(4.32-4.35)	4.36	(4.34-4.37)
Median	5.75	(5.74-5.77)	5.77	(5.76-5.79)
0.75Q	7.62	(7.61-7.64)	7.65	(7.64-7.67)
0.90Q	9.83	(9.81-9.84)	9.87	(9.85-9.88)
0.95Q	11.44	(11.43-11.46)	11.49	(11.47-11.50)
0.99Q	15.23	(15.20-15.26)	15.28	(15.25-15.31)
MSE $\hat{\mu}$	0.00	-	0.00	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.10	-	0.09	-

8.3.2 With alternative numbers for the data generation method

Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.61	(4.53-4.69)	4.63	(4.55-4.71)
$\hat{\beta}$	0.78	(0.77-0.79)	0.78	(0.77-0.79)
$\hat{\pi}$	0.46	(0.45-0.46)	0.45	(0.44-0.46)
Mean	5.84	(5.81-5.88)	5.84	(5.81-5.88)
0.25Q	3.82	(3.79-3.86)	3.85	(3.81-3.88)
Median	5.41	(5.37-5.44)	5.43	(5.40-5.47)
0.75Q	7.39	(7.36-7.42)	7.42	(7.39-7.45)
0.90Q	9.54	(9.51-9.57)	9.58	(9.54-9.61)
0.95Q	11.01	(10.97-11.04)	11.04	(11.01-11.08)
0.99Q	14.12	(14.08-14.16)	14.16	(14.12-14.20)
MSE $\hat{\alpha}$	1.89	-	1.87	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.23	-	0.22	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.53	(4.47-4.59)	4.53	(4.46-4.59)
$\hat{\beta}$	0.77	(0.76-0.78)	0.77	(0.76-0.78)
$\hat{\pi}$	0.45	(0.44-0.46)	0.45	(0.44-0.45)
Mean	5.82	(5.80-5.85)	5.82	(5.80-5.85)
0.25Q	3.80	(3.77-3.83)	3.81	(3.79-3.84)
Median	5.39	(5.36-5.42)	5.40	(5.38-5.43)
0.75Q	7.37	(7.35-7.40)	7.40	(7.37-7.42)
0.90Q	9.53	(9.50-9.55)	9.55	(9.53-9.58)
0.95Q	10.99	(10.97-11.01)	11.02	(11.00-11.05)
0.99Q	14.10	(14.07-14.13)	14.15	(14.11-14.18)
MSE $\hat{\alpha}$	1.26	-	1.25	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.21	-	0.21	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.47	(4.41-4.53)	4.46	(4.41-4.52)
$\hat{\beta}$	0.76	(0.76-0.77)	0.76	(0.75-0.77)
$\hat{\pi}$	0.45	(0.44-0.46)	0.45	(0.44-0.45)
Mean	5.82	(5.80-5.85)	5.82	(5.80-5.85)
0.25Q	3.79	(3.77-3.82)	3.80	(3.78-3.83)
Median	5.39	(5.36-5.41)	5.40	(5.38-5.43)
0.75Q	7.38	(7.36-7.41)	7.40	(7.38-7.43)
0.90Q	9.55	(9.52-9.57)	9.57	(9.55-9.60)
0.95Q	11.02	(10.99-11.04)	11.05	(11.03-11.08)
0.99Q	14.15	(14.12-14.18)	14.19	(14.16-14.22)
MSE $\hat{\alpha}$	1.12	-	1.13	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.21	-	0.21	-

Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.63	(4.56-4.69)	4.63	(4.56-4.69)
$\hat{\beta}$	0.78	(0.77-0.79)	0.77	(0.77-0.78)
$\hat{\pi}$	0.50	(0.49-0.51)	0.49	(0.49-0.50)
Mean	5.91	(5.88-5.94)	5.91	(5.88-5.94)
0.25Q	3.88	(3.85-3.91)	3.89	(3.86-3.92)
Median	5.48	(5.45-5.51)	5.50	(5.47-5.52)
0.75Q	7.47	(7.44-7.50)	7.49	(7.47-7.52)
0.90Q	9.63	(9.60-9.65)	9.66	(9.63-9.68)
0.95Q	11.09	(11.06-11.12)	11.13	(11.10-11.15)
0.99Q	14.21	(14.17-14.24)	14.25	(14.21-14.29)
MSE $\hat{\alpha}$	1.19	-	1.18	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.17	-	0.17	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.58	(4.53-4.64)	4.58	(4.53-4.63)
$\hat{\beta}$	0.77	(0.76-0.78)	0.77	(0.76-0.77)
$\hat{\pi}$	0.49	(0.49-0.50)	0.49	(0.48-0.49)
Mean	5.92	(5.90-5.94)	5.92	(5.90-5.94)
0.25Q	3.88	(3.86-3.91)	3.89	(3.87-3.92)
Median	5.49	(5.46-5.51)	5.5	(5.48-5.53)
0.75Q	7.49	(7.46-7.51)	7.51	(7.48-7.53)
0.90Q	9.65	(9.63-9.67)	9.68	(9.66-9.70)
0.95Q	11.12	(11.1-11.14)	11.16	(11.13-11.18)
0.99Q	14.25	(14.21-14.28)	14.29	(14.26-14.32)
MSE $\hat{\alpha}$	0.91	-	0.92	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.16	-	0.16	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.55	(4.51-4.60)	4.54	(4.50-4.59)
$\hat{\beta}$	0.77	(0.76-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.49	(0.49-0.50)	0.49	(0.49-0.50)
Mean	5.91	(5.89-5.93)	5.91	(5.89-5.93)
0.25Q	3.88	(3.85-3.90)	3.88	(3.86-3.90)
Median	5.48	(5.46-5.50)	5.49	(5.47-5.51)
0.75Q	7.48	(7.46-7.50)	7.50	(7.48-7.52)
0.90Q	9.65	(9.63-9.67)	9.68	(9.66-9.70)
0.95Q	11.12	(11.10-11.14)	11.15	(11.13-11.17)
0.99Q	14.25	(14.22-14.27)	14.29	(14.27-14.32)
MSE $\hat{\alpha}$	0.75	-	0.76	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.16	-	0.16	-

Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.70	(4.64-4.76)	4.69	(4.63-4.75)
$\hat{\beta}$	0.78	(0.78-0.79)	0.78	(0.77-0.79)
$\hat{\pi}$	0.55	(0.54-0.55)	0.54	(0.54-0.55)
Mean	5.98	(5.95-6.00)	5.98	(5.95-6.00)
0.25Q	3.94	(3.92-3.97)	3.95	(3.93-3.98)
Median	5.55	(5.52-5.57)	5.56	(5.54-5.59)
0.75Q	7.54	(7.52-7.57)	7.56	(7.54-7.59)
0.90Q	9.70	(9.68-9.72)	9.73	(9.71-9.75)
0.95Q	11.16	(11.14-11.19)	11.2	(11.17-11.23)
0.99Q	14.27	(14.23-14.31)	14.32	(14.28-14.36)
MSE $\hat{\alpha}$	0.98	-	0.98	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.13	-	0.13	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.66	(4.62-4.71)	4.65	(4.61-4.70)
$\hat{\beta}$	0.78	(0.77-0.78)	0.77	(0.77-0.78)
$\hat{\pi}$	0.55	(0.54-0.55)	0.54	(0.54-0.55)
Mean	5.98	(5.96-6.00)	5.98	(5.96-6.00)
0.25Q	3.94	(3.92-3.97)	3.95	(3.93-3.97)
Median	5.55	(5.53-5.57)	5.56	(5.54-5.58)
0.75Q	7.55	(7.53-7.57)	7.57	(7.55- 7.59)
0.90Q	9.71	(9.69-9.73)	9.74	(9.72-9.76)
0.95Q	11.17	(11.16-11.19)	11.21	(11.19- 11.23)
0.99Q	14.28	(14.26-14.31)	14.33	(14.3-14.36)
MSE $\hat{\alpha}$	0.66	-	0.67	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.13	-	0.12	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.60	(4.56-4.64)	4.59	(4.55-4.62)
$\hat{\beta}$	0.77	(0.76-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.55	(0.54-0.55)	0.54	(0.54-0.55)
Mean	5.97	(5.96-5.99)	5.97	(5.96-5.99)
0.25Q	3.93	(3.91-3.95)	3.94	(3.92-3.95)
Median	5.54	(5.52-5.56)	5.55	(5.54-5.57)
0.75Q	7.55	(7.53-7.56)	7.57	(7.55-7.58)
0.90Q	9.72	(9.70-9.73)	9.74	(9.73-9.76)
0.95Q	11.19	(11.17-11.20)	11.22	(11.21-11.24)
0.99Q	14.31	(14.29-14.33)	14.36	(14.34-14.38)
MSE $\hat{\alpha}$	0.51	-	0.53	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.12	-	0.12	-

8.3.3 Incubation period estimates when π is removed

Gamma	Deng's method		Qin's method	
$m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	8.77	(8.69-8.84)	8.74	(8.67-8.82)
$\hat{\beta}$	1.11	(1.10-1.12)	1.10	(1.10-1.11)
Mean	7.90	(7.88- 7.91)	7.90	(7.88-7.91)
0.25Q	5.96	(5.94-5.98)	5.96	(5.94-5.98)
Median	7.59	(7.57-7.61)	7.6	(7.58-7.62)
0.75Q	9.50	(9.48-9.52)	9.51	(9.49-9.53)
0.90Q	11.47	(11.45-11.50)	11.49	(11.47-11.51)
0.95Q	12.77	(12.74-12.80)	12.79	(12.76-12.82)
0.99Q	15.45	(15.41-15.49)	15.48	(15.44-15.52)
MSE $\hat{\alpha}$	15.73	-	15.56	-
MSE $\hat{\beta}$	0.12	-	0.11	-
Gamma	Deng's method		Qin's method	
$m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	8.83	(8.76-8.9)	8.8	(8.73-8.87)
$\hat{\beta}$	1.12	(1.11-1.12)	1.11	(1.10- 1.12)
Mean	7.91	(7.89-7.92)	7.91	(7.89-7.92)
0.25Q	5.98	(5.96-6.00)	5.98	(5.96-6.00)
Median	7.61	(7.59-7.62)	7.61	(7.60-7.63)
0.75Q	9.51	(9.50-9.53)	9.52	(9.51-9.54)
0.90Q	11.47	(11.45-11.49)	11.49	(11.47-11.51)
0.95Q	12.76	(12.74-12.78)	12.78	(12.76- 12.80)
0.99Q	15.43	(15.39-15.46)	15.45	(15.42-15.49)
MSE $\hat{\alpha}$	15.82	-	15.60	-
MSE $\hat{\beta}$	0.11	-	0.11	-
Gamma	Deng's method		Qin's method	
$m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	8.78	(8.71-8.84)	8.75	(8.68-8.82)
$\hat{\beta}$	1.11	(1.10-1.12)	1.1	(1.10-1.11)
Mean	7.91	(7.89-7.92)	7.91	(7.89-7.92)
0.25Q	5.97	(5.96-5.99)	5.97	(5.96-5.99)
Median	7.60	(7.59-7.62)	7.61	(7.60-7.63)
0.75Q	9.51	(9.50-9.53)	9.52	(9.51-9.54)
0.90Q	11.48	(11.46-11.50)	11.5	(11.48-11.51)
0.95Q	12.77	(12.75-12.79)	12.79	(12.77-12.81)
0.99Q	15.44	(15.42-15.47)	15.47	(15.44-15.50)
MSE $\hat{\alpha}$	15.41	-	15.19	-
MSE $\hat{\beta}$	0.11	-	0.11	-

Weibull	Deng's method		Qin's method	
$m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.73	(2.71-2.75)	2.72	(2.70-2.74)
$\hat{k}$	10.34	(10.31-10.37)	10.35	(10.32-10.38)
Mean	9.21	(9.18-9.23)	9.21	(9.18-9.23)
0.25Q	6.53	(6.49-6.57)	6.53	(6.49-6.57)
Median	9.03	(9.00-9.06)	9.03	(9.00-9.07)
0.75Q	11.67	(11.64-11.70)	11.68	(11.65-11.71)
0.90Q	14.09	(14.06-14.12)	14.11	(14.08-14.13)
0.95Q	15.54	(15.51-15.57)	15.56	(15.52-15.59)
0.99Q	18.24	(18.19-18.29)	18.27	(18.21-18.32)
MSE $\hat{\lambda}$	0.63	-	0.63	-
MSE $\hat{k}$	5.70	-	5.73	-
Weibull	Deng's method		Qin's method	
$m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.71	(2.70-2.73)	2.71	(2.69-2.72)
$\hat{k}$	10.32	(10.30-10.35)	10.33	(10.31-10.35)
Mean	9.19	(9.16-9.21)	9.19	(9.16-9.21)
0.25Q	6.51	(6.48-6.54)	6.51	(6.48 - 6.54)
Median	9.01	(8.98-9.04)	9.01	(8.99-9.04)
0.75Q	11.65	(11.63-11.68)	11.66	(11.64-11.68)
0.90Q	14.07	(14.05-14.10)	14.09	(14.07-14.11)
0.95Q	15.52	(15.50-15.55)	15.54	(15.52-15.57)
0.99Q	18.22	(18.18-18.26)	18.25	(18.21-18.29)
MSE $\hat{\lambda}$	0.57	-	0.57	-
MSE $\hat{k}$	5.54	-	5.57	-
Weibull	Deng's method		Qin's method	
$m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.70	(2.68-2.71)	2.70	(2.68-2.71)
$\hat{k}$	10.31	(10.29-10.33)	10.31	(10.29-10.34)
Mean	9.17	(9.15-9.19)	9.17	(9.15-9.19)
0.25Q	6.49	(6.46-6.51)	6.49	(6.46-6.51)
Median	8.99	(8.97-9.02)	9.00	(8.97-9.02)
0.75Q	11.64	(11.62-11.66)	11.65	(11.63-11.67)
0.90Q	14.07	(14.05-14.09)	14.08	(14.06-14.10)
0.95Q	15.52	(15.50-15.55)	15.54	(15.51-15.56)
0.99Q	18.23	(18.19-18.27)	18.25	(18.21-18.29)
MSE $\hat{\lambda}$	0.54	-	0.54	-
MSE $\hat{k}$	5.45	-	5.47	-

Log-normal	Deng's method		Qin's method	
$m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	2.05	(2.05-2.05)	2.05	(2.05-2.05)
$\hat{\sigma}$	0.31	(0.31-0.32)	0.32	(0.31-0.32)
Mean	8.15	(8.13-8.17)	8.15	(8.13-8.17)
0.25Q	6.28	(6.25-6.3)	6.28	(6.26-6.30)
Median	7.76	(7.74-7.77)	7.76	(7.74-7.78)
0.75Q	9.59	(9.57-9.60)	9.60	(9.58-9.61)
0.90Q	11.61	(11.58-11.63)	11.62	(11.60-11.64)
0.95Q	13.01	(12.98-13.04)	13.03	(13.00- 13.06)
0.99Q	16.14	(16.09-16.19)	16.17	(16.12-16.22)
MSE $\hat{\mu}$	0.06	-	0.06	-
MSE $\hat{\sigma}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	2.05	(2.04-2.05)	2.05	(2.04-2.05)
$\hat{\sigma}$	0.32	(0.31-0.32)	0.32	(0.31-0.32)
Mean	8.13	(8.12-8.15)	8.13	(8.12-8.15)
0.25Q	6.26	(6.24-6.27)	6.26	(6.24-6.28)
Median	7.74	(7.72-7.75)	7.74	(7.72-7.76)
0.75Q	9.57	(9.55-9.58)	9.58	(9.56-9.59)
0.90Q	11.59	(11.57-11.61)	11.60	(11.58-11.62)
0.95Q	13.00	(12.97-13.02)	13.02	(12.99-13.04)
0.99Q	16.12	(16.08-16.16)	16.15	(16.11-16.19)
MSE $\hat{\mu}$	0.06	-	0.06	-
MSE $\hat{\sigma}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	2.05	(2.04-2.05)	2.05	(2.04-2.05)
$\hat{\sigma}$	0.32	(0.32-0.32)	0.32	(0.32-0.32)
Mean	8.14	(8.13-8.15)	8.14	(8.13-8.15)
0.25Q	6.25	(6.23-6.27)	6.25	(6.24-6.27)
Median	7.74	(7.72-7.75)	7.74	(7.73-7.76)
0.75Q	9.58	(9.57-9.60)	9.59	(9.58-9.60)
0.90Q	11.62	(11.60-11.63)	11.63	(11.61-11.65)
0.95Q	13.04	(13.01-13.06)	13.06	(13.03-13.08)
0.99Q	16.19	(16.15-16.23)	16.22	(16.18-16.26)
MSE $\hat{\mu}$	0.06	-	0.06	-
MSE $\hat{\sigma}$	0.01	-	0.01	-

8.3.4 Incubation period estimates Deng and Qin

Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.45	(4.37-4.52)	4.49	(4.42-4.56)
$\hat{\beta}$	0.76	(0.75-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.11	(0.10-0.12)	0.10	(0.09-0.11)
Mean	5.76	(5.72-5.80)	5.76	(5.72-5.80)
0.25Q	3.74	(3.70-3.78)	3.78	(3.75-3.82)
Median	5.32	(5.27-5.36)	5.37	(5.33-5.41)
0.75Q	7.31	(7.27-7.35)	7.36	(7.33-7.40)
0.90Q	9.48	(9.44-9.51)	9.53	(9.49-9.56)
0.95Q	10.95	(10.91-10.99)	11.00	(10.97-11.04)
0.99Q	14.10	(14.06-14.14)	14.14	(14.10-14.18)
MSE $\hat{\alpha}$	1.76	-	1.51	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.03	-	0.02	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.48	(4.43-4.54)	4.49	(4.44-4.55)
$\hat{\beta}$	0.76	(0.75-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.09	(0.08-0.09)	0.08	(0.08-0.09)
Mean	5.84	(5.81-5.87)	5.84	(5.81-5.87)
0.25Q	3.81	(3.77-3.84)	3.83	(3.80-3.86)
Median	5.40	(5.37-5.44)	5.42	(5.39-5.46)
0.75Q	7.40	(7.36-7.43)	7.42	(7.39-7.45)
0.90Q	9.56	(9.53-9.59)	9.59	(9.56-9.62)
0.95Q	11.03	(11.00-11.07)	11.06	(11.03-11.09)
0.99Q	14.17	(14.14-14.20)	14.20	(14.17-14.23)
MSE $\hat{\alpha}$	1.11	-	1.01	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.02	-	0.02	-
Gamma	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	4.53	(4.48-4.58)	4.52	(4.47-4.57)
$\hat{\beta}$	0.76	(0.76-0.77)	0.76	(0.76-0.77)
$\hat{\pi}$	0.08	(0.07-0.08)	0.07	(0.07-0.08)
Mean	5.89	(5.86-5.92)	5.89	(5.86-5.92)
0.25Q	3.86	(3.83-3.88)	3.86	(3.83-3.89)
Median	5.45	(5.42-5.48)	5.46	(5.43-5.49)
0.75Q	7.45	(7.42-7.48)	7.46	(7.43-7.49)
0.90Q	9.61	(9.59-9.64)	9.63	(9.60-9.66)
0.95Q	11.08	(11.06-11.11)	11.10	(11.07-11.13)
0.99Q	14.21	(14.18-14.24)	14.23	(14.20-14.26)
MSE $\hat{\alpha}$	0.83	-	0.80	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.01	-	0.01	-

Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.01	(4.92-5.10)	5.02	(4.93-5.11)
$\hat{\beta}$	0.80	(0.79-0.81)	0.80	(0.79-0.81)
$\hat{\pi}$	0.14	(0.13-0.14)	0.13	(0.12-0.14)
Mean	6.18	(6.13-6.22)	6.18	(6.13-6.22)
0.25Q	4.13	(4.08-4.18)	4.15	(4.11-4.20)
Median	5.75	(5.70-5.80)	5.78	(5.73-5.82)
0.75Q	7.76	(7.71-7.81)	7.79	(7.75-7.83)
0.90Q	9.93	(9.88-9.97)	9.96	(9.92-10.00)
0.95Q	11.40	(11.35-11.44)	11.43	(11.39-11.47)
0.99Q	14.51	(14.46-14.55)	14.54	(14.50-14.59)
MSE $\hat{\alpha}$	2.06	-	1.95	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.02	-	0.02	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.01	(4.94-5.08)	5.01	(4.94-5.08)
$\hat{\beta}$	0.80	(0.79-0.80)	0.79	(0.79-0.80)
$\hat{\pi}$	0.12	(0.11-0.13)	0.12	(0.11-0.12)
Mean	6.22	(6.18-6.26)	6.22	(6.18-6.26)
0.25Q	4.17	(4.14-4.21)	4.18	(4.15-4.22)
Median	5.80	(5.76-5.84)	5.81	(5.77-5.85)
0.75Q	7.81	(7.77-7.85)	7.83	(7.79-7.87)
0.90Q	9.98	(9.94-10.01)	10.00	(9.96-10.03)
0.95Q	11.44	(11.40-11.47)	11.47	(11.43-11.50)
0.99Q	14.54	(14.50-14.57)	14.57	(14.54-14.61)
MSE $\hat{\alpha}$	1.32	-	1.30	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Gamma	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.07	(5.01-5.13)	5.05	(4.99-5.11)
$\hat{\beta}$	0.80	(0.80-0.81)	0.80	(0.79-0.80)
$\hat{\pi}$	0.11	(0.10-0.11)	0.11	(0.10-0.11)
Mean	6.27	(6.24-6.30)	6.27	(6.24-6.30)
0.25Q	4.23	(4.19- 4.26)	4.23	(4.20-4.26)
Median	5.86	(5.82- 5.89)	5.86	(5.83-5.90)
0.75Q	7.87	(7.83-7.90)	7.88	(7.84-7.91)
0.90Q	10.02	(9.99-10.06)	10.04	(10.01-10.08)
0.95Q	11.48	(11.45-11.51)	11.51	(11.47-11.54)
0.99Q	14.57	(14.53-14.60)	14.60	(14.57-14.63)
MSE $\hat{\alpha}$	0.94	-	0.96	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.01	-	0.01	-

Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.42	(5.32-5.52)	5.43	(5.33-5.52)
$\hat{\beta}$	0.83	(0.82-0.84)	0.83	(0.82-0.84)
$\hat{\pi}$	0.19	(0.18-0.20)	0.19	(0.18-0.20)
Mean	6.39	(6.35-6.44)	6.39	(6.35-6.44)
0.25Q	4.36	(4.31-4.41)	4.37	(4.33-4.42)
Median	5.99	(5.94-6.03)	6.00	(5.96-6.05)
0.75Q	7.99	(7.94-8.03)	8.01	(7.96-8.06)
0.90Q	10.13	(10.08-10.17)	10.16	(10.11-10.20)
0.95Q	11.57	(11.53-11.62)	11.61	(11.56-11.65)
0.99Q	14.63	(14.58-14.67)	14.66	(14.62-14.71)
MSE $\hat{\alpha}$	2.73	-	2.72	-
MSE $\hat{\beta}$	0.03	-	0.03	-
MSE $\hat{\pi}$	0.02	-	0.02	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.31	(5.23-5.39)	5.30	(5.22-5.38)
$\hat{\beta}$	0.82	(0.82-0.83)	0.82	(0.81-0.83)
$\hat{\pi}$	0.19	(0.18-0.20)	0.19	(0.18-0.19)
Mean	6.37	(6.34-6.41)	6.37	(6.34-6.41)
0.25Q	4.33	(4.30-4.37)	4.34	(4.30-4.38)
Median	5.96	(5.92-6.00)	5.98	(5.94-6.02)
0.75Q	7.97	(7.93-8.00)	7.99	(7.95-8.02)
0.90Q	10.11	(10.08-10.15)	10.14	(10.10-10.17)
0.95Q	11.56	(11.52-11.59)	11.59	(11.55-11.62)
0.99Q	14.61	(14.58-14.65)	14.65	(14.62-14.69)
MSE $\hat{\alpha}$	1.72	-	1.74	-
MSE $\hat{\beta}$	0.02	-	0.02	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Gamma	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\alpha}$	5.28	(5.21-5.34)	5.27	(5.20-5.34)
$\hat{\beta}$	0.82	(0.81-0.83)	0.82	(0.81-0.82)
$\hat{\pi}$	0.19	(0.18-0.19)	0.18	(0.18-0.19)
Mean	6.37	(6.34-6.40)	6.37	(6.34-6.40)
0.25Q	4.33	(4.30-4.37)	4.34	(4.31-4.38)
Median	5.96	(5.93-6.00)	5.98	(5.94-6.01)
0.75Q	7.97	(7.93-8.00)	7.99	(7.95-8.02)
0.90Q	10.11	(10.08-10.14)	10.14	(10.10-10.17)
0.95Q	11.56	(11.53-11.59)	11.59	(11.56-11.62)
0.99Q	14.61	(14.58-14.64)	14.65	(14.62-14.68)
MSE $\hat{\alpha}$	0.57	-	0.59	-
MSE $\hat{\beta}$	0.01	-	0.01	-
MSE $\hat{\pi}$	0.11	-	0.11	-

Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	1.90	(1.89-1.92)	1.99	(1.98-2.01)
$\hat{k}$	7.42	(7.36-7.48)	7.87	(7.84-7.90)
$\hat{\pi}$	0.13	(0.11-0.14)	0.02	(0.02-0.03)
Mean	6.60	(6.55-6.65)	6.60	(6.55-6.65)
0.25Q	3.86	(3.81-3.91)	4.21	(4.18-4.24)
Median	6.12	(6.06-6.17)	6.55	(6.52-6.58)
0.75Q	8.82	(8.76-8.87)	9.28	(9.25-9.31)
0.90Q	11.53	(11.48-11.59)	11.99	(11.96-12.02)
0.95Q	13.27	(13.21-13.32)	13.69	(13.66-13.73)
0.99Q	16.68	(16.63-16.74)	17.02	(16.98-17.07)
MSE $\hat{\lambda}$	0.07	-	0.04	-
MSE $\hat{k}$	1.16	-	0.25	-
MSE $\hat{\pi}$	0.05	-	0.00	-
Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	1.92	(1.91-1.94)	2.00	(1.99-2.01)
$\hat{k}$	7.57	(7.53-7.62)	7.96	(7.94-7.98)
$\hat{\pi}$	0.09	(0.08-0.09)	0.01	(0.01-0.01)
Mean	6.73	(6.69-6.76)	6.73	(6.69-6.76)
0.25Q	3.97	(3.93-4.01)	4.27	(4.25-4.29)
Median	6.26	(6.22-6.30)	6.63	(6.61-6.65)
0.75Q	8.98	(8.93-9.02)	9.38	(9.36-9.39)
0.90Q	11.70	(11.65-11.74)	12.09	(12.07-12.11)
0.95Q	13.42	(13.38-13.46)	13.80	(13.77-13.82)
0.99Q	16.81	(16.77-16.85)	17.12	(17.09-17.15)
MSE $\hat{\lambda}$	0.04	-	0.02	-
MSE $\hat{k}$	0.68	-	0.11	-
MSE $\hat{\pi}$	0.02	-	0.00	-
Weibull	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	1.93	(1.92-1.93)	2.00	(1.99-2.01)
$\hat{k}$	7.59	(7.56-7.63)	7.97	(7.96-7.99)
$\hat{\pi}$	0.08	(0.07-0.08)	0.01	(0.00-0.01)
Mean	6.74	(6.71-6.78)	6.74	(6.71-6.78)
0.25Q	3.98	(3.95-4.01)	4.28	(4.26-4.29)
Median	6.28	(6.24-6.32)	6.64	(6.62- 6.65)
0.75Q	9.00	(8.96-9.04)	9.39	(9.37-9.41)
0.90Q	9.61	(9.59-9.64)	9.63	(9.60-9.66)
0.95Q	13.45	(13.41-13.48)	13.82	(13.8-13.83)
0.99Q	16.83	(16.79-16.87)	17.14	(17.12-17.17)
MSE $\hat{\lambda}$	0.03	-	0.01	-
MSE $\hat{k}$	0.56	-	0.07	-
MSE $\hat{\pi}$	0.02	-	0.00	

Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.00	(1.98-2.02)	2.10	(2.09-2.12)
$\hat{k}$	7.91	(7.86-7.97)	8.39	(8.36-8.43)
$\hat{\pi}$	0.14	(0.13-0.15)	0.04	(0.04-0.05)
Mean	7.03	(6.98-7.08)	7.03	(6.98-7.08)
0.25Q	4.25	(4.20-4.30)	4.64	(4.61-4.67)
Median	6.59	(6.53-6.65)	7.05	(7.01-7.08)
0.75Q	9.32	(9.26-9.38)	9.81	(9.78-9.84)
0.90Q	12.04	(11.98-12.09)	12.5	(12.47-12.54)
0.95Q	13.75	(13.70-13.80)	14.18	(14.14-14.22)
0.99Q	17.10	(17.04-17.15)	17.43	(17.38-17.48)
MSE $\hat{\lambda}$	0.07	-	0.05	-
MSE $\hat{k}$	0.90	-	0.46	-
MSE $\hat{\pi}$	0.03	-	0.01	-
Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.01	(2.00-2.02)	2.11	(2.10-2.12)
$\hat{k}$	8.01	(7.96-8.06)	8.47	(8.44-8.50)
$\hat{\pi}$	0.12	(0.11-0.13)	0.03	(0.03-0.04)
Mean	7.10	(7.06-7.15)	7.10	(7.06-7.15)
0.25Q	4.31	(4.27-4.36)	4.69	(4.66-4.71)
Median	6.67	(6.62-6.72)	7.12	(7.09-7.14)
0.75Q	9.42	(9.37-9.47)	9.89	(9.87-9.92)
0.90Q	12.15	(12.10-12.19)	12.6	(12.57-12.62)
0.95Q	13.86	(13.81-13.90)	14.28	(14.25-14.31)
0.99Q	17.2	(17.15-17.24)	17.53	(17.49-17.56)
MSE $\hat{\lambda}$	0.04	-	0.03	-
MSE $\hat{k}$	0.66	-	0.40	-
MSE $\hat{\pi}$	0.02	-	0.01	-
Weibull	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.01	(2.00-2.02)	2.11	(2.10-2.12)
$\hat{k}$	8.03	(7.98-8.08)	8.50	(8.47-8.52)
$\hat{\pi}$	0.11	(0.10-0.12)	0.03	(0.02-0.03)
Mean	7.12	(7.08-7.16)	7.12	(7.08-7.16)
0.25Q	4.33	(4.29-4.37)	4.71	(4.69-4.73)
Median	6.69	(6.65-6.74)	7.14	(7.12-7.16)
0.75Q	9.45	(9.40-9.49)	9.92	(9.90-9.94)
0.90Q	12.17	(12.12-12.21)	12.63	(12.60-12.65)
0.95Q	13.88	(13.84-13.92)	14.31	(14.28-14.33)
0.99Q	17.21	(17.17-17.25)	17.55	(17.52-17.58)
MSE $\hat{\lambda}$	0.03	-	0.02	-
MSE $\hat{k}$	0.55	-	0.38	-
MSE $\hat{\pi}$	0.02	-	0.01	-

Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.06	(2.04-2.08)	2.19	(2.17-2.20)
$\hat{k}$	8.17	(8.11-8.24)	8.69	(8.65-8.73)
$\hat{\pi}$	0.20	(0.19-0.21)	0.10	(0.09-0.11)
Mean	7.25	(7.20-7.31)	7.25	(7.20-7.31)
0.25Q	4.47	(4.42-4.53)	4.91	(4.87-4.95)
Median	6.84	(6.78-6.90)	7.34	(7.30-7.39)
0.75Q	9.58	(9.52-9.64)	10.10	(10.05-10.14)
0.90Q	12.28	(12.22-12.34)	12.75	(12.71-12.80)
0.95Q	13.97	(13.91-14.03)	14.40	(14.35-14.45)
0.99Q	17.26	(17.21-17.32)	17.56	(17.51-17.62)
MSE $\hat{\lambda}$	0.09	-	0.08	-
MSE $\hat{k}$	1.08	-	0.97	-
MSE $\hat{\pi}$	0.03	-	0.02	-
Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.05	(2.04-2.07)	2.18	(2.17-2.19)
$\hat{k}$	8.19	(8.14-8.25)	8.74	(8.70-8.77)
$\hat{\pi}$	0.19	(0.18-0.19)	0.09	(0.08-0.09)
Mean	7.27	(7.22-7.32)	7.27	(7.22-7.32)
0.25Q	4.47	(4.42-4.52)	4.93	(4.90-4.96)
Median	6.85	(6.80-6.91)	7.38	(7.35-7.42)
0.75Q	9.61	(9.56-9.67)	10.15	(10.12-10.19)
0.90Q	12.32	(12.27-12.37)	12.82	(12.79-12.86)
0.95Q	14.02	(13.97-14.07)	14.48	(14.44-14.52)
0.99Q	17.32	(17.28-17.37)	17.65	(17.61-17.70)
MSE $\hat{\lambda}$	0.05	-	0.06	-
MSE $\hat{k}$	0.81	-	0.88	-
MSE $\hat{\pi}$	0.02	-	0.02	-
Weibull	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\lambda}$	2.04	(2.03-2.06)	2.18	(2.17-2.19)
$\hat{k}$	8.17	(8.12-8.22)	8.74	(8.71-8.77)
$\hat{\pi}$	0.19	(0.18-0.19)	0.09	(0.08 0.09)
Mean	7.24	(7.20-7.29)	7.24	(7.20-7.29)
0.25Q	4.45	(4.40-4.49)	4.94	(4.91-4.96)
Median	6.83	(6.78-6.87)	7.39	(7.36-7.42)
0.75Q	9.58	(9.54-9.63)	10.16	(10.12-10.19)
0.90Q	12.29	(12.25-12.34)	12.82	(12.79-12.86)
0.95Q	13.99	(13.95-14.03)	14.47	(14.44-14.51)
0.99Q	17.29	(17.25-17.33)	17.64	(17.6-17.68)
MSE $\hat{\lambda}$	0.04	-	0.05	-
MSE $\hat{k}$	0.64	-	0.83	-
MSE $\hat{\pi}$	0.02	-	0.02	-

Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.73	(1.73-1.74)	1.74	(1.73-1.74)
$\hat{\sigma}$	0.42	(0.42-0.42)	0.42	(0.42-0.42)
$\hat{\pi}$	0.07	(0.07-0.08)	0.07	(0.06-0.07)
Mean	6.20	(6.17-6.23)	6.20	(6.17-6.23)
0.25Q	4.29	(4.26-4.32)	4.32	(4.29-4.35)
Median	5.68	(5.65-5.71)	5.71	(5.68-5.75)
0.75Q	7.52	(7.49-7.56)	7.56	(7.53-7.59)
0.90Q	9.70	(9.66-9.73)	9.74	(9.70-9.77)
0.95Q	11.29	(11.25-11.32)	11.33	(11.30-11.37)
0.99Q	15.02	(14.98-15.07)	15.07	(15.02-15.11)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.75	(1.74-1.75)	1.75	(1.75-1.76)
$\hat{\sigma}$	0.42	(0.41-0.42)	0.41	(0.41-0.42)
$\hat{\pi}$	0.06	(0.05-0.06)	0.05	(0.05-0.06)
Mean	6.28	(6.25-6.30)	6.28	(6.25-6.30)
0.25Q	4.36	(4.33-4.38)	4.39	(4.36-4.41)
Median	5.76	(5.73-5.79)	5.79	(5.76-5.82)
0.75Q	7.61	(7.58-7.64)	7.65	(7.62-7.68)
0.90Q	9.79	(9.76-9.82)	9.83	(9.81-9.86)
0.95Q	11.38	(11.35-11.41)	11.43	(11.40-11.45)
0.99Q	15.11	(15.07-15.14)	15.15	(15.12-15.19)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.76	(1.75-1.76)	1.76	(1.76-1.76)
$\hat{\sigma}$	0.41	(0.41-0.41)	0.41	(0.41-0.41)
$\hat{\pi}$	0.05	(0.04-0.05)	0.04	(0.04-0.05)
Mean	6.31	(6.29-6.34)	6.31	(6.29-6.34)
0.25Q	4.39	(4.37-4.41)	4.42	(4.40-4.44)
Median	5.80	(5.77-5.82)	5.83	(5.81-5.85)
0.75Q	7.66	(7.63-7.68)	7.69	(7.67-7.71)
0.90Q	9.83	(9.81-9.86)	9.87	(9.85-9.90)
0.95Q	11.42	(11.40-11.45)	11.47	(11.44-11.49)
0.99Q	15.14	(15.11-15.17)	15.18	(15.16-15.21)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-

Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.79	(1.79-1.80)	1.80	(1.79-1.80)
$\hat{\sigma}$	0.40	(0.40-0.41)	0.40	(0.40-0.41)
$\hat{\pi}$	0.11	(0.11-0.12)	0.11	(0.10-0.12)
Mean	6.54	(6.51-6.58)	6.54	(6.51-6.58)
0.25Q	4.61	(4.57-4.65)	4.64	(4.60-4.67)
Median	6.03	(5.99-6.07)	6.07	(6.03-6.10)
0.75Q	7.90	(7.86-7.94)	7.94	(7.90-7.98)
0.90Q	10.08	(10.05-10.12)	10.13	(10.09-10.16)
0.95Q	11.67	(11.64-11.71)	11.72	(11.68-11.75)
0.99Q	15.37	(15.32-15.41)	15.41	(15.37-15.46)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.80	(1.80-1.81)	1.81	(1.80-1.81)
$\hat{\sigma}$	0.40	(0.40-0.40)	0.40	(0.40-0.40)
$\hat{\pi}$	0.10	(0.10-0.11)	0.10	(0.09-0.10)
Mean	6.59	(6.56-6.62)	6.59	(6.56-6.62)
0.25Q	4.65	(4.62-4.68)	4.68	(4.65-4.71)
Median	6.08	(6.05-6.11)	6.12	(6.09-6.15)
0.75Q	7.96	(7.93-7.99)	8.00	(7.97-8.03)
0.90Q	10.14	(10.11-10.17)	10.19	(10.15-10.22)
0.95Q	11.73	(11.69-11.76)	11.77	(11.74-11.80)
0.99Q	15.4	(15.37-15.44)	15.45	(15.41-15.49)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.1$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.81	(1.80-1.81)	1.81	(1.81-1.82)
$\hat{\sigma}$	0.40	(0.40-0.40)	0.40	(0.40-0.40)
$\hat{\pi}$	0.09	(0.09-0.10)	0.09	(0.09-0.09)
Mean	6.61	(6.59-6.64)	6.61	(6.59-6.64)
0.25Q	4.67	(4.65-4.70)	4.71	(4.68-4.73)
Median	6.11	(6.08-6.14)	6.15	(6.12-6.17)
0.75Q	7.99	(7.96-8.01)	8.03	(8.00-8.06)
0.90Q	10.17	(10.14-10.19)	10.21	(10.18-10.24)
0.95Q	11.75	(11.72-11.78)	11.80	(11.77-11.82)
0.99Q	15.42	(15.39-15.44)	15.47	(15.44-15.49)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-

Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 600$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.82	(1.81-1.82)	1.82	(1.82-1.83)
$\hat{\sigma}$	0.39	(0.39-0.40)	0.39	(0.39-0.40)
$\hat{\pi}$	0.19	(0.18-0.19)	0.18	(0.17-0.19)
Mean	6.69	(6.65-6.73)	6.69	(6.65-6.73)
0.25Q	4.77	(4.73-4.80)	4.80	(4.76-4.83)
Median	6.20	(6.16-6.24)	6.23	(6.19-6.27)
0.75Q	8.06	(8.02-8.10)	8.10	(8.06-8.14)
0.90Q	10.23	(10.19-10.26)	10.27	(10.23-10.31)
0.95Q	11.79	(11.76-11.83)	11.84	(11.80-11.88)
0.99Q	15.42	(15.37-15.47)	15.47	(15.42-15.52)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1200$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.81	(1.81-1.82)	1.82	(1.81-1.82)
$\hat{\sigma}$	0.40	(0.39-0.40)	0.40	(0.39-0.40)
$\hat{\pi}$	0.19	(0.18-0.19)	0.18	(0.18-0.19)
Mean	6.65	(6.62-6.67)	6.65	(6.62-6.67)
0.25Q	4.71	(4.68-4.74)	4.74	(4.71-4.77)
Median	6.15	(6.12-6.18)	6.18	(6.15-6.21)
0.75Q	8.02	(7.99-8.05)	8.06	(8.03-8.09)
0.90Q	10.19	(10.16-10.22)	10.24	(10.21-10.26)
0.95Q	11.77	(11.74-11.79)	11.81	(11.78-11.84)
0.99Q	15.41	(15.38-15.45)	15.46	(15.43-15.49)
MSE $\hat{\mu}$	0.01	-	0.01	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.01	-	0.01	-
Log-normal	Deng's method		Qin's method	
$\pi = 0.2$, $m = 1800$	Estimate	(95%CI)	Estimate	(95%CI)
$\hat{\mu}$	1.81	(1.80-1.81)	1.81	(1.81-1.82)
$\hat{\sigma}$	0.40	(0.40-0.40)	0.40	(0.39-0.40)
$\hat{\pi}$	0.19	(0.19-0.20)	0.19	(0.18-0.19)
Mean	6.61	(6.59-6.63)	6.61	(6.59-6.63)
0.25Q	4.68	(4.66-4.70)	4.71	(4.69-4.73)
Median	6.11	(6.09-6.13)	6.14	(6.12-6.17)
0.75Q	7.98	(7.96-8.00)	8.02	(8.00-8.05)
0.90Q	10.16	(10.13-10.18)	10.20	(10.18-10.22)
0.95Q	11.73	(11.71-11.75)	11.78	(11.76-11.80)
0.99Q	15.38	(15.35-15.41)	15.43	(15.40-15.46)
MSE $\hat{\mu}$	0.00	-	0.00	-
MSE $\hat{\sigma}$	0.00	-	0.00	-
MSE $\hat{\pi}$	0.00	-	0.00	-

8.4 R-Code

8.4.1 Data generation outbreak Wuhan

```
simulate_outbreak_wuhan <- function(total_cases, leaveenter_per_day_
    prior19th, leave_per_day_from19thonwards, growth_rate, distribution,
     par1, par2, time){

  infected_list <- list()
  incubation_periods <- list()
  infected_travel <- list()
  time_obs <- list()
  left_wuhan <- list()
  population_total <- c(1:11000000) #population Wuhan
  population_minus_infected <- c(1:11000000) # same as population_total
      at the start - to ensure that not > 1 infections per person

  day_of_inf <- vector(mode = "list", length = 19)

  for (i in 1:19){
    #for infected_list- draw a sample from people who are still in
       Wuhan and have not been infected before -> population_minus_
       infected
    infected_list[[i]] <- sample(x = population_minus_infected, size =
       total_cases, replace = FALSE)
    if (i < 15){
      #add people who enter city
      population_total <- c(population_total, (tail(population_total,
         1) + 1):(tail(population_total, 1) + leaveenter_per_day_
         prior19th))
      #for left_wuhan - draw a sample from everyone that is in the city
          on that day
      left_wuhan[[i]] <- sample(x = population_total, size = leaveenter
         _per_day_prior19th, replace = FALSE)
      population_total <- population_total[!(population_total %in% left
         _wuhan[[i]])]
      population_minus_infected <- population_total[!(population_total
         %in% unlist(infected_list[1:i]))]
    }
    else{
      #draw people who travel out Wuhan from population
      left_wuhan[[i]] <- sample(x = population_total, size = leave_per_
         day_from19thonwards, replace = FALSE)
      population_total <- population_total[!(population_total %in% left
         _wuhan[[i]])]
      population_minus_infected <- population_total[!(population_total
         %in% unlist(infected_list[1:i]))]
      infected_travel[[i]] <- intersect(left_wuhan[[i]], unlist(
         infected_list))

      #look up time of infection for each infected traveler in List of
         infected
      for (k in 1:length(infected_travel[[i]])){
        day_of_inf[[i]][k] <- which(sapply(X = infected_list, FUN =
          function(X) infected_travel[[i]][k] %in% X))[1]
      }
      #for those the compute incubation periods are
      if(distribution == "gamma"){
```

```
        incubation_periods[[i]] <- rgamma(n = length(infected_travel[[i
            ]]), shape = par1, rate = par2)
      } else if(distribution == "weibull"){
        incubation_periods[[i]] <- rweibull(n = length(infected_travel
            [[i]]), shape = par1, scale = par2)
      } else if(distribution == "lognormal"){
        incubation_periods[[i]] <- rlnorm(n = length(infected_travel[[i
            ]]), meanlog = par1, sdlog = par2)
      }

      if(time == "forward"){ #forward time
        #the forward times when symptom onset is not before day of
            travel are:
        time_obs[[i]] <- (day_of_inf[[i]][day_of_inf[[i]] + incubation_
            periods[[i]] > i] +
                            incubation_periods[[i]][day_of_inf[[i]] +
                                incubation_periods[[i]] > i]) - i
      }
      else if(time == "backward"){ #backward time
        time_obs[[i]] <- i - (day_of_inf[[i]][day_of_inf[[i]] +
            incubation_periods[[i]] > i])
      }
    }
    #number of cases per day
    total_cases <- total_cases + round(total_cases * (exp(growth_rate)
        - 1))
  }
  return(unlist(time_obs))
}
```

8.4.2 Simulation study incubation period estimates

```
incubationPeriod_ParEstimates <- function(runs, total_cases, leaveenter
    _per_day_prior19th, leave_per_day_from19thonwards, growth_rate,
    distribution, par1, par2){

  result = NULL

  aa <- par1
  bb <- par2


  if(distribution == "gamma"){
    ########################################gamma incubation
    f <- function(y,a,b){
      return(dgamma(y,shape=a,rate=b))
    }

    #pdf of forward time
    h <- function(y,a,b){
      RE = b/a*(1-pgamma(y,shape=a,rate=b))
      return(RE)
    }
    FI <- function(y,a,b) return(pgamma(y,shape=a,rate=b))

    F <- function(k,a,b,p){
```

```
    RE = rep(0,length(k))
    posit = which(k>0)
    REh = pgamma(k[posit],shape=a+1,rate=b)+k[posit]*b/a*(1-pgamma(k[
        posit],shape=a,rate=b))
    if (p>0.999) RE[posit] = REh
    else{
      REf = pgamma(k[posit],shape=a,rate=b)
      RE[posit] = REf*(1-p)+REh*p
    }
    return(RE)
  }


}else if(distribution == "weibull"){
  ##########################################weibull incubation
  f <- function(y,a,b){
    return(dweibull(y,shape=a,scale=b))
  }
  h <- function(y,a,b){
    RE = (1-pweibull(y,shape=a,scale=b))/gamma(1+1/a)/b
    return(RE)
  }
  FI <- function(y,a,b) return(pweibull(y,shape=a,scale=b))
  F <- function(k,a,b,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REh = pgamma((k[posit]/b)^a,shape=1/a,rate=1)
    if (p>0.999) RE[posit] = REh
    else{
      REf = pweibull(k[posit],shape=a,scale=b)
      RE[posit] = REf*(1-p)+REh*p
    }
    return(RE)
  }

}else if(distribution == "lognormal"){
  ##########################################lognormal incubation
  f <- function(y,u,s){
    RE = rep(0,length(y))
    posit = which(y>0)
    RE[posit]=dnorm(log(y[posit]),u,s)/y[posit]
    return(RE)
  }
  h <- function(y,u,s){
    RE = exp(-u-s^2/2)*(1-pnorm(log(y),u,s))
    return(RE)
  }
  FI <- function(y,u,s){
    RE = rep(0,length(y))
    posit = which(y>0)
    RE[posit]=pnorm(log(y[posit]),u,s)
    return(RE)
  }
  F <- function(k,u,s,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REf = pnorm(log(k[posit]),u,s)
```

```
    REh = pnorm(log(k[posit]),u+s^2,s)+exp(-u-s^2/2)*k[posit]*(1-
        pnorm(log(k[posit]),u,s))
    RE[posit] = REf*(1-p)+REh*p  # equation 2 of paper Deng
    return(RE)
  }

}

loglik <- function(pa){
  a = pa[1]
  b = pa[2]
  p = pa[3]
  if (p>0.999) P=h(x,a,b)
  else P = p*h(x,a,b)+(1-p)*f(x,a,b)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}


Loglik <- function(pa){
  a = pa[1]
  b = pa[2]
  p = pa[3]
  P = F(x+0.5,a,b,p)-F(x-0.5,a,b,p)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}



for (w in 1:9){
  if (w%%3==1) m=600
  if (w%%3==2) m=1200
  if (w%%3==0) m=1800
  if (w<=3) pp=1
  if (w>=4 & w <=6) pp=0.9
  if (w>=7) pp=0.8


  a.vec = numeric(runs)
  b.vec = numeric(runs)
  q.vec = numeric(runs)
  ave.vec = numeric(runs)
  Q1.vec = numeric(runs)
  Q2.vec = numeric(runs)
  Q3.vec = numeric(runs)
  Q4.vec = numeric(runs)
  Q5.vec = numeric(runs)
  Q6.vec = numeric(runs)

  a.vec.qin = numeric(runs)
  b.vec.qin = numeric(runs)
  q.vec.qin = numeric(runs)
  ave.vec.qin = numeric(runs)
```

```
Q1.vec.qin = numeric(runs)
Q2.vec.qin = numeric(runs)
Q3.vec.qin = numeric(runs)
Q4.vec.qin = numeric(runs)
Q5.vec.qin = numeric(runs)
Q6.vec.qin = numeric(runs)

bias_par1_Deng = numeric(runs)
bias_par2_Deng = numeric(runs)
bias_p_Deng = numeric(runs)

bias_par1_Qin = numeric(runs)
bias_par2_Qin = numeric(runs)
bias_p_Qin = numeric(runs)


for (run in 1:runs){
  data_forwardtime <- simulate_outbreak_wuhan(total_cases,
     leaveenter_per_day_prior19th, leave_per_day_from19thonwards,
     growth_rate, distribution, par1, par2, "forward")
  #### Generating data
  set.seed(run)
  x = rep(NA,m)
  for (i in 1:m){
    select = rbinom(1,1,pp) #additional infected at day of travel
    if (select==0){
      if(distribution == "gamma"){
        x[i] = rgamma(1,aa,bb)
      }else if(distribution == "weibull"){
        x[i] = rweibull(1,aa,bb)
      }else if(distribution == "lognormal"){
        x[i] = rlnorm(1,aa,bb)
      }
    }
    else{#infected before day of travel
      x[i] = sample(x = data_forwardtime, size = 1) #observed
         forward time
    }
  }

  #### p: censor probability

  x = round(x,0)
  x[x>25] = 25



  if(distribution == "gamma"){
    par_1 <- optim(par=c(4,0.5,0.8),Loglik,method='L-BFGS-B',
                   lower=c(1.1,0.1,0),upper=c(10,2,1))$par
  }
  else if(distribution == "weibull"){
    par_1 <- optim(par=c(2,10,0.8),Loglik,method='L-BFGS-B',
                   lower=c(1.1,2,0),upper=c(5,15,1))$par

  }
  else if(distribution == "lognormal"){
    par_1 <- optim(par=c(2,0.4,0.8),Loglik,method='L-BFGS-B',
```

```
                    lower=c(1,0.1,0),upper=c(5,1,1))$par
}

#model parameter and pi

a = par_1[1]
b = par_1[2]
p = par_1[3]

a.vec[run] = a
b.vec[run] = b
q.vec[run] = 1-p



#mean, median, quartiles
if(distribution == "gamma"){
  ##########################gamma
  est <- c(
    c(a,b,1-p) ,  # estimates for parameters alpha and beta, and
       pi
    a/b,          # mean
    round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
       estimated quantiles
  # -round(loglik(c(a,b)),2),

} else if(distribution == "weibull"){
  ##########################weibull
  est <- c(
    c(a,b,1-p) ,   # estimates for parameters k and lambda, and
       pi
    b*gamma(1+1/a) , # mean
    round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
       ),2))  # estimated quantiles
  #-round(loglik(c(a,b)),2),

} else if(distribution == "lognormal"){
  ##########################lognormal
  est <- c(
    c(a,b,1-p),  # estimates for parameters mu and sigma, and pi
    exp(a+b^2/2), # mean
    round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
       estimated quantiles
  #-round(loglik(c(a,b)),2)

}

#mean, median, quartiles, bias

ave.vec[run] = est[4]
Q1.vec[run] = est[5]
Q2.vec[run] = est[6]
Q3.vec[run] = est[7]
Q4.vec[run] = est[8]
Q5.vec[run] = est[9]
Q6.vec[run] = est[10]
```

```
bias_par1_Deng[run] <- est[1] - aa
bias_par2_Deng[run] <- est[2] - bb
bias_p_Deng[run] <- est[3] - (1 - pp)



if(distribution == "gamma"){
  par_2 <- optim(par=c(4,0.5,0.8),loglik,method='L-BFGS-B',
                 lower=c(1.1,0.1,0),upper=c(10,2,1))$par
}
else if(distribution == "weibull"){
  par_2 <- optim(par=c(2,10,0.8),loglik,method='L-BFGS-B',
                 lower=c(1.1,2,0),upper=c(5,15,1))$par
}
else if(distribution == "lognormal"){
  par_2 <- optim(par=c(2,0.4,0.8),loglik,method='L-BFGS-B',
                 lower=c(1,0.1,0),upper=c(5,1,1))$par
}

#model parameter and pi

a = par_2[1]
b = par_2[2]
p = par_2[3]

a.vec.qin[run] = a
b.vec.qin[run] = b
q.vec.qin[run] = 1-p


#mean, median, quartiles
if(distribution == "gamma"){
  ##########################gamma
  est <- c(
    c(a,b,1-p),  # estimates for parameters alpha and beta, and
       pi
    a/b,         # mean
    round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
       estimated quantiles
  # -round(loglik(c(a,b)),2),

} else if(distribution == "weibull"){
  ##########################weibull
  est <- c(
    c(a,b,1-p),   # estimates for parameters k and lambda, and pi
    b*gamma(1+1/a), # mean
    round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
       ),2))  # estimated quantiles
  #-round(loglik(c(a,b)),2),

} else if(distribution == "lognormal"){
  ##########################lognormal
  est <- c(
    c(a,b,1-p),  # estimates for parameters mu and sigma, and pi
    exp(a+b^2/2), # mean
    round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
       estimated quantiles
```

```
    #-round(loglik(c(a,b)),2)

  }


  #mean, median, quartiles, bias

  ave.vec.qin[run] = est[4]
  Q1.vec.qin[run] = est[5]
  Q2.vec.qin[run] = est[6]
  Q3.vec.qin[run] = est[7]
  Q4.vec.qin[run] = est[8]
  Q5.vec.qin[run] = est[9]
  Q6.vec.qin[run] = est[10]


  bias_par1_Qin[run] <- est[1] - aa
  bias_par2_Qin[run] <- est[2] - bb
  bias_p_Qin[run] <- est[3] - (1 - pp)


}

#MSE and coverage
MSE_par1_Deng <- sum(bias_par1_Deng^2)/runs
MSE_par2_Deng <- sum(bias_par2_Deng^2)/runs
MSE_p_Deng <- sum(bias_p_Deng^2)/runs

MSE_par1_Qin <- sum(bias_par1_Qin^2)/runs
MSE_par2_Qin <- sum(bias_par2_Qin^2)/runs
MSE_p_Qin <- sum(bias_p_Qin^2)/runs



res <- rbind(c(round(mean(a.vec),2), round(t.test(a.vec)$conf.int
    [1], 2), round(t.test(a.vec)$conf.int[2], 2)),
             c(round(mean(b.vec),2), round(t.test(b.vec)$conf.int
                [1], 2), round(t.test(b.vec)$conf.int[2], 2)),
             c(round(mean(q.vec),2), round(t.test(q.vec)$conf.int
                [1], 2), round(t.test(q.vec)$conf.int[2], 2)),
             c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
             c(round(mean(Q1.vec),2), round(t.test(Q1.vec)$conf.int
                [1], 2), round(t.test(Q1.vec)$conf.int[2], 2)),
             c(round(mean(Q2.vec),2), round(t.test(Q2.vec)$conf.int
                [1], 2), round(t.test(Q2.vec)$conf.int[2], 2)),
             c(round(mean(Q3.vec),2), round(t.test(Q3.vec)$conf.int
                [1], 2), round(t.test(Q3.vec)$conf.int[2], 2)),
             c(round(mean(Q4.vec),2), round(t.test(Q4.vec)$conf.int
                [1], 2), round(t.test(Q4.vec)$conf.int[2], 2)),
             c(round(mean(Q5.vec),2), round(t.test(Q5.vec)$conf.int
                [1], 2), round(t.test(Q5.vec)$conf.int[2], 2)),
             c(round(mean(Q6.vec),2), round(t.test(Q6.vec)$conf.int
                [1], 2), round(t.test(Q6.vec)$conf.int[2], 2)),
             c(round(MSE_par1_Deng, 2), NULL, NULL),
             c(round(MSE_par2_Deng, 2), NULL, NULL),
             c(round(MSE_p_Deng, 2), NULL, NULL),
```

```
                    c(round(mean(a.vec.qin),2), round(t.test(a.vec.qin)$
                       conf.int[1], 2), round(t.test(a.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(b.vec.qin),2), round(t.test(b.vec.qin)$
                       conf.int[1], 2), round(t.test(b.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(q.vec.qin),2), round(t.test(q.vec.qin)$
                       conf.int[1], 2), round(t.test(q.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                       int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
                    c(round(mean(Q1.vec.qin),2), round(t.test(Q1.vec.qin)$
                       conf.int[1], 2), round(t.test(Q1.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(Q2.vec.qin),2), round(t.test(Q2.vec.qin)$
                       conf.int[1], 2), round(t.test(Q2.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(Q3.vec.qin),2), round(t.test(Q3.vec.qin)$
                       conf.int[1], 2), round(t.test(Q3.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(Q4.vec.qin),2), round(t.test(Q4.vec.qin)$
                       conf.int[1], 2), round(t.test(Q4.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(Q5.vec.qin),2), round(t.test(Q5.vec.qin)$
                       conf.int[1], 2), round(t.test(Q5.vec.qin)$conf.int
                       [2], 2)),
                    c(round(mean(Q6.vec.qin),2), round(t.test(Q6.vec.qin)$
                       conf.int[1], 2), round(t.test(Q6.vec.qin)$conf.int
                       [2], 2)),
                    c(round(MSE_par1_Qin, 2), NULL, NULL),
                    c(round(MSE_par2_Qin, 2), NULL, NULL),
                    c(round(MSE_p_Qin, 2), NULL, NULL))



    parameter <- c('alpha','beta', 'pi', 'mean', '.25Q', 'median', '.75
       Q', '.90Q', '.95Q', '.99Q', 'MSE␣alpha', 'MSE␣beta', 'MSE␣p',
                   'alpha','beta', 'pi', 'mean', '.25Q', 'median', '.75
                      Q', '.90Q', '.95Q', '.99Q', 'MSE␣alpha', 'MSE␣
                      beta', 'MSE␣p')
    method <- c(rep('Deng', nrow(res) / 2), rep('Qin', nrow(res) / 2))

    colnames(res) <- c("estimate", "L_95%CI", "U_95%CI")

    res <- data.frame(parameter, method, res)

    result = rbind(result, res)
    print(cat("sample␣size:", m, "pi:", 1-pp))
    print(res)
  }
  return(result)
}
```

8.4.3 Simulation study incubation period estimates when π is removed

```
Estimates_WihtoutPi <- function(runs, total_cases, leaveenter_per_day_
```

```
 prior19th, leave_per_day_from19thonwards, growth_rate, distribution,
  par1, par2){

result = NULL

aa <- par1
bb <- par2
pp <- 1


if(distribution == "gamma"){
  ########################################gamma incubation
  f <- function(y,a,b){
    return(dgamma(y,shape=a,rate=b))
  }
  h <- function(y,a,b){
    RE = b/a*(1-pgamma(y,shape=a,rate=b))
    return(RE)
  }
  FI <- function(y,a,b) return(pgamma(y,shape=a,rate=b))
  F <- function(k,a,b,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REh = pgamma(k[posit],shape=a+1,rate=b)+k[posit]*b/a*(1-pgamma(k[
       posit],shape=a,rate=b))
    if (p>0.999) RE[posit] = REh
    else{
      REf = pgamma(k[posit],shape=a,rate=b)
      RE[posit] = REf*(1-p)+REh*p
    }
    return(RE)
  }



}else if(distribution == "weibull"){
  ########################################weibull incubation
  f <- function(y,a,b){
    return(dweibull(y,shape=a,scale=b))
  }
  h <- function(y,a,b){
    RE = (1-pweibull(y,shape=a,scale=b))/gamma(1+1/a)/b
    return(RE)
  }
  FI <- function(y,a,b) return(pweibull(y,shape=a,scale=b))
  F <- function(k,a,b,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REh = pgamma((k[posit]/b)^a,shape=1/a,rate=1)
    if (p>0.999) RE[posit] = REh
    else{
      REf = pweibull(k[posit],shape=a,scale=b)
      RE[posit] = REf*(1-p)+REh*p
    }
    return(RE)
  }

}else if(distribution == "lognormal"){
```

```
  ##########################################lognormal incubation
  f <- function(y,u,s){
    RE = rep(0,length(y))
    posit = which(y>0)
    RE[posit]=dnorm(log(y[posit]),u,s)/y[posit]
    return(RE)
  }
  h <- function(y,u,s){
    RE = exp(-u-s^2/2)*(1-pnorm(log(y),u,s))
    return(RE)
  }
  FI <- function(y,u,s){
    RE = rep(0,length(y))
    posit = which(y>0)
    RE[posit]=pnorm(log(y[posit]),u,s)
    return(RE)
  }
  F <- function(k,u,s,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REh = pnorm(log(k[posit]),u+s^2,s)+exp(-u-s^2/2)*k[posit]*(1-
       pnorm(log(k[posit]),u,s))
    if (p>0.999) RE[posit]=REh
    else {
      REf = pnorm(log(k[posit]),u,s)
      RE[posit] = REf*(1-p)+REh*p
    }
    return(RE)
  }
}


loglik0 <- function(pa){
  a = pa[1]
  b = pa[2]
  P = h(x,a,b)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}


Loglik0 <- function(pa){
  a = pa[1]
  b = pa[2]
  P = F(x+0.5,a,b,1)-F(x-0.5,a,b,1)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}



for (w in 1:3){
  if (w%%3==1) m=600
  if (w%%3==2) m=1200
```

```
if (w%%3==0) m=1800



a.vec = numeric(runs)
b.vec = numeric(runs)
ave.vec = numeric(runs)
Q1.vec = numeric(runs)
Q2.vec = numeric(runs)
Q3.vec = numeric(runs)
Q4.vec = numeric(runs)
Q5.vec = numeric(runs)
Q6.vec = numeric(runs)

a.vec.qin = numeric(runs)
b.vec.qin = numeric(runs)
ave.vec.qin = numeric(runs)
Q1.vec.qin = numeric(runs)
Q2.vec.qin = numeric(runs)
Q3.vec.qin = numeric(runs)
Q4.vec.qin = numeric(runs)
Q5.vec.qin = numeric(runs)
Q6.vec.qin = numeric(runs)

bias_par1_Deng = numeric(runs)
bias_par2_Deng = numeric(runs)


bias_par1_Qin = numeric(runs)
bias_par2_Qin = numeric(runs)




for (run in 1:runs){

  data_forwardtime <- simulate_outbreak_wuhan(total_cases,
     leaveenter_per_day_prior19th, leave_per_day_from19thonwards,
     growth_rate, distribution, par1, par2, "forward")

  #### Generating data
  #set.seed(run)
  x = rep(NA,m)
  for (i in 1:m){
    x[i] = sample(x = data_forwardtime, size = 1)   #observed
       forward time
  }



  #### p: censor probability
  x = round(x,0)
  x[x>25] = 25
```

```
if(distribution == "gamma"){
  par_1 <- optim(par=c(4,0.5),Loglik0,method='L-BFGS-B',
                 lower=c(1.1,0.1),upper=c(10,2))$par
}
else if(distribution == "weibull"){
  par_1 <- optim(par=c(2,10),Loglik0,method='L-BFGS-B',
                 lower=c(1.1,2),upper=c(5,15))$par

}
else if(distribution == "lognormal"){
  par_1 <- optim(par=c(2,0.4),Loglik0,method='L-BFGS-B',
                 lower=c(1,0.1),upper=c(5,1))$par
}

#model parameter

a = par_1[1]
b = par_1[2]


a.vec[run] = a
b.vec[run] = b


#mean, median, quartiles

if(distribution == "gamma"){
  ##########################gamma
  est <- c(
    c(a,b) ,  # estimates for parameters alpha and beta, and pi
    a/b,          # mean
    round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
       estimated quantiles
  # -round(loglik(c(a,b)),2),

} else if(distribution == "weibull"){
  ##########################weibull
  est <- c(
    c(a,b) ,   # estimates for parameters k and lambda, and pi
    b*gamma(1+1/a) , # mean
    round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
       ),2))  # estimated quantiles
  #-round(loglik(c(a,b)),2),

} else if(distribution == "lognormal"){
  ##########################lognormal
  est <- c(
    c(a,b),  # estimates for parameters mu and sigma, and pi
    exp(a+b^2/2), # mean
    round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
       estimated quantiles
  #-round(loglik(c(a,b)),2)
}

ave.vec[run] = est[3]
Q1.vec[run] = est[4]
Q2.vec[run] = est[5]
Q3.vec[run] = est[6]
```

```
Q4.vec[run] = est[7]
Q5.vec[run] = est[8]
Q6.vec[run] = est[9]


bias_par1_Deng[run] <- est[1] - aa
bias_par2_Deng[run] <- est[2] - bb



if(distribution == "gamma"){
  par_2 <- optim(par=c(4,0.5,0.8),loglik0,method='L-BFGS-B',
                 lower=c(1.1,0.1,0),upper=c(10,2,1))$par
}
else if(distribution == "weibull"){
  par_2 <- optim(par=c(2,10,0.8),loglik0,method='L-BFGS-B',
                 lower=c(1.1,2,0),upper=c(5,15,1))$par
}
else if(distribution == "lognormal"){
  par_2 <- optim(par=c(2,0.4,0.8),loglik0,method='L-BFGS-B',
                 lower=c(1,0.1,0),upper=c(5,1,1))$par
}

#model parameter and pi

a = par_2[1]
b = par_2[2]


a.vec.qin[run] = a
b.vec.qin[run] = b




#mean, median, quartiles
if(distribution == "gamma"){
  ##########################gamma
  est <- c(
    c(a,b),  # estimates for parameters alpha and beta, and pi
    a/b,          # mean
    round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
       estimated quantiles
  # -round(loglik(c(a,b)),2),

} else if(distribution == "weibull"){
  ##########################weibull
  est <- c(
    c(a,b),   # estimates for parameters k and lambda, and pi
    b*gamma(1+1/a), # mean
    round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
       ),2))  # estimated quantiles
  #-round(loglik(c(a,b)),2),

} else if(distribution == "lognormal"){
  ##########################lognormal
```

```
    est <- c(
      c(a,b),  # estimates for parameters mu and sigma, and pi
      exp(a+b^2/2), # mean
      round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
         estimated quantiles
    #-round(loglik(c(a,b)),2)
  }


  #mean, median, quantiles
  ave.vec.qin[run] = est[3]
  Q1.vec.qin[run] = est[4]
  Q2.vec.qin[run] = est[5]
  Q3.vec.qin[run] = est[6]
  Q4.vec.qin[run] = est[7]
  Q5.vec.qin[run] = est[8]
  Q6.vec.qin[run] = est[9]


  bias_par1_Qin[run] <- est[1] - aa
  bias_par2_Qin[run] <- est[2] - bb

}



#MSE and coverage
MSE_par1_Deng <- sum(bias_par1_Deng^2)/runs
MSE_par2_Deng <- sum(bias_par2_Deng^2)/runs


MSE_par1_Qin <- sum(bias_par1_Qin^2)/runs
MSE_par2_Qin <- sum(bias_par2_Qin^2)/runs



res <- rbind(c(round(mean(a.vec),2), round(t.test(a.vec)$conf.int
    [1], 2), round(t.test(a.vec)$conf.int[2], 2)),
             c(round(mean(b.vec),2), round(t.test(b.vec)$conf.int
                [1], 2), round(t.test(b.vec)$conf.int[2], 2)),
             c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
             c(round(mean(Q1.vec),2), round(t.test(Q1.vec)$conf.int
                [1], 2), round(t.test(Q1.vec)$conf.int[2], 2)),
             c(round(mean(Q2.vec),2), round(t.test(Q2.vec)$conf.int
                [1], 2), round(t.test(Q2.vec)$conf.int[2], 2)),
             c(round(mean(Q3.vec),2), round(t.test(Q3.vec)$conf.int
                [1], 2), round(t.test(Q3.vec)$conf.int[2], 2)),
             c(round(mean(Q4.vec),2), round(t.test(Q4.vec)$conf.int
                [1], 2), round(t.test(Q4.vec)$conf.int[2], 2)),
             c(round(mean(Q5.vec),2), round(t.test(Q5.vec)$conf.int
                [1], 2), round(t.test(Q5.vec)$conf.int[2], 2)),
             c(round(mean(Q6.vec),2), round(t.test(Q6.vec)$conf.int
                [1], 2), round(t.test(Q6.vec)$conf.int[2], 2)),
             c(round(MSE_par1_Deng, 2), NULL, NULL),
             c(round(MSE_par2_Deng, 2), NULL, NULL),
```

```
                    c(round(mean(a.vec.qin),2), round(t.test(a.vec.qin)$
                        conf.int[1], 2), round(t.test(a.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(b.vec.qin),2), round(t.test(b.vec.qin)$
                        conf.int[1], 2), round(t.test(b.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                        int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
                    c(round(mean(Q1.vec.qin),2), round(t.test(Q1.vec.qin)$
                        conf.int[1], 2), round(t.test(Q1.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(Q2.vec.qin),2), round(t.test(Q2.vec.qin)$
                        conf.int[1], 2), round(t.test(Q2.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(Q3.vec.qin),2), round(t.test(Q3.vec.qin)$
                        conf.int[1], 2), round(t.test(Q3.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(Q4.vec.qin),2), round(t.test(Q4.vec.qin)$
                        conf.int[1], 2), round(t.test(Q4.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(Q5.vec.qin),2), round(t.test(Q5.vec.qin)$
                        conf.int[1], 2), round(t.test(Q5.vec.qin)$conf.int
                        [2], 2)),
                    c(round(mean(Q6.vec.qin),2), round(t.test(Q6.vec.qin)$
                        conf.int[1], 2), round(t.test(Q6.vec.qin)$conf.int
                        [2], 2)),
                    c(round(MSE_par1_Qin, 2), NULL, NULL),
                    c(round(MSE_par2_Qin, 2), NULL, NULL))


    parameter <- c('alpha','beta','mean', '.25Q', 'median', '.75Q', '
        .90Q', '.95Q', '.99Q', 'MSE␣alpha', 'MSE␣beta',
                    'alpha','beta', 'mean', '.25Q', 'median', '.75Q', '
                        .90Q', '.95Q', '0.99Q', 'MSE␣alpha', 'MSE␣beta')
    method <- c(rep('Deng', nrow(res) / 2), rep('Qin', nrow(res) / 2))

    colnames(res) <- c("estimate", "L_95%CI", "U_95%CI")

    res <- data.frame(parameter, method, res)

    result = rbind(result, res)
    print(cat("sample␣size:", m))
    print(res)
  }
  return(result)
}
```

8.4.4 Data generation Deng and Qin

```
GenerateData_Deng <- function(m, pi, par1, par2, distribution, time){

  pp <- 1 - pi
  for (run in 1:1000){
```

```
    #### Generating data
    set.seed(run)
    obs_time = rep(NA,m)
    for (i in 1:m){
      select = rbinom(1,1,pp)
      if (select==0){
        if(distribution == "gamma"){
          obs_time[i] = rgamma(1,par1,par2)
        }
        else if(distribution == "weibull"){
          obs_time[i] = rweibull(1,par1,par2)
        }
        else if(distribution == "lognormal"){
          obs_time[i] = exp(rnorm(1,par1,par2))
        }
      }
      else{
        while (TRUE){
          C = runif(1,0,30)              #departure time
          if(distribution == "gamma"){
            Y = rgamma(1,par1,par2)
          }
          else if(distribution == "weibull"){
            Y = rweibull(1,par1,par2)
          }
          else if(distribution == "lognormal"){
            Y = exp(rnorm(1,par1,par2))
          }
          if (Y>C) break
        }
        if(time == "forward"){
          obs_time[i] <- round(Y-C,4) #observed forward time
        }
        else if(time == "backward"){
          obs_time[i] = round(Y-(Y-C),4)                #observed
             backward time
        }
      }
    }
  }
  return(obs_time)
}
```

8.4.5 Simulation study incubation period estimates Deng and Qin

```
incubationPeriod_Estimates_Deng <- function(runs, distribution, par1,
   par2){

  result = NULL

  aa <- par1
  bb <- par2


  if(distribution == "gamma"){
```

```
    ########################################gamma incubation
    f <- function(y,a,b){
      return(dgamma(y,shape=a,rate=b))
    }

    #pdf of forward time
    h <- function(y,a,b){
      RE = b/a*(1-pgamma(y,shape=a,rate=b))
      return(RE)
    }
    FI <- function(y,a,b) return(pgamma(y,shape=a,rate=b))

    F <- function(k,a,b,p){
      RE = rep(0,length(k))
      posit = which(k>0)
      REh = pgamma(k[posit],shape=a+1,rate=b)+k[posit]*b/a*(1-pgamma(k[
          posit],shape=a,rate=b))
      if (p>0.999) RE[posit] = REh
      else{
        REf = pgamma(k[posit],shape=a,rate=b)
        RE[posit] = REf*(1-p)+REh*p
      }
      return(RE)
    }



  }else if(distribution == "weibull"){
    ########################################weibull incubation
    f <- function(y,a,b){
      return(dweibull(y,shape=a,scale=b))
    }
    h <- function(y,a,b){
      RE = (1-pweibull(y,shape=a,scale=b))/gamma(1+1/a)/b
      return(RE)
    }
    FI <- function(y,a,b) return(pweibull(y,shape=a,scale=b))
    F <- function(k,a,b,p){
      RE = rep(0,length(k))
      posit = which(k>0)
      REh = pgamma((k[posit]/b)^a,shape=1/a,rate=1)
      if (p>0.999) RE[posit] = REh
      else{
        REf = pweibull(k[posit],shape=a,scale=b)
        RE[posit] = REf*(1-p)+REh*p
      }
      return(RE)
    }

  }else if(distribution == "lognormal"){
    ########################################lognormal incubation
    f <- function(y,u,s){
      RE = rep(0,length(y))
      posit = which(y>0)
      RE[posit]=dnorm(log(y[posit]),u,s)/y[posit]
      return(RE)
    }
    h <- function(y,u,s){
```

```
    RE = exp(-u-s^2/2)*(1-pnorm(log(y),u,s))
    return(RE)
  }
  FI <- function(y,u,s){
    RE = rep(0,length(y))
    posit = which(y>0)
    RE[posit]=pnorm(log(y[posit]),u,s)
    return(RE)
  }
  F <- function(k,u,s,p){
    RE = rep(0,length(k))
    posit = which(k>0)
    REf = pnorm(log(k[posit]),u,s)
    REh = pnorm(log(k[posit]),u+s^2,s)+exp(-u-s^2/2)*k[posit]*(1-
       pnorm(log(k[posit]),u,s))
    RE[posit] = REf*(1-p)+REh*p  # equation 2 of paper Deng
    return(RE)
  }

}

loglik <- function(pa){
  a = pa[1]
  b = pa[2]
  p = pa[3]
  if (p>0.999) P=h(x,a,b)
  else P = p*h(x,a,b)+(1-p)*f(x,a,b)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}


Loglik <- function(pa){
  a = pa[1]
  b = pa[2]
  p = pa[3]
  P = F(x+0.5,a,b,p)-F(x-0.5,a,b,p)
  P[P<0.00001]=0.00001
  RE = - sum( log(P) )
  if (RE>10000) RE=10000
  return(RE)
}



for (w in 1:9){
  if (w%%3==1) m=600
  if (w%%3==2) m=1200
  if (w%%3==0) m=1800
  if (w<=3) pp=1
  if (w>=4 & w <=6) pp=0.9
  if (w>=7) pp=0.8


  a.vec = numeric(runs)
  b.vec = numeric(runs)
```

```
q.vec = numeric(runs)
ave.vec = numeric(runs)
Q1.vec = numeric(runs)
Q2.vec = numeric(runs)
Q3.vec = numeric(runs)
Q4.vec = numeric(runs)
Q5.vec = numeric(runs)
Q6.vec = numeric(runs)

a.vec.qin = numeric(runs)
b.vec.qin = numeric(runs)
q.vec.qin = numeric(runs)
ave.vec.qin = numeric(runs)
Q1.vec.qin = numeric(runs)
Q2.vec.qin = numeric(runs)
Q3.vec.qin = numeric(runs)
Q4.vec.qin = numeric(runs)
Q5.vec.qin = numeric(runs)
Q6.vec.qin = numeric(runs)

bias_par1_Deng = numeric(runs)
bias_par2_Deng = numeric(runs)
bias_p_Deng = numeric(runs)

bias_par1_Qin = numeric(runs)
bias_par2_Qin = numeric(runs)
bias_p_Qin = numeric(runs)


for (run in 1:runs){
  #### Generating data
  set.seed(run)
  x = rep(NA,m)
  for (i in 1:m){
    select = rbinom(1,1,pp) #additional infected at day of travel
    if (select==0){
      if(distribution == "gamma"){
        x[i] = rgamma(1,aa,bb)
      }else if(distribution == "weibull"){
        x[i] = rweibull(1,aa,bb)
      }else if(distribution == "lognormal"){
        x[i] = rlnorm(1,aa,bb)
      }
    }
    else{
      while (TRUE){
        C = runif(1,0,30)           #departure time
        if(distribution == "gamma"){
          Y = rgamma(1,par1,par2)
        }
        else if(distribution == "weibull"){
          Y = rweibull(1,par1,par2)
        }
        else if(distribution == "lognormal"){
          Y = exp(rnorm(1,par1,par2))
        }
        if (Y>C) break
      }
```

```
    x[i] <- (Y-C) #observed forward time
  }
}
#(suggestion for correction: select <- size 1000 0/1, x-1000 from
    gamma (if gamma),y-1000 simulate_outbreakWuhan, x*select+y*
   (1-select), ignore for now)

#### p: censor probability

x = round(x,0)
x[x>25] = 25



if(distribution == "gamma"){
  par_1 <- optim(par=c(4,0.5,0.8),Loglik,method='L-BFGS-B',
                 lower=c(1.1,0.1,0),upper=c(10,2,1))$par
}
else if(distribution == "weibull"){
  par_1 <- optim(par=c(2,10,0.8),Loglik,method='L-BFGS-B',
                 lower=c(1.1,2,0),upper=c(5,15,1))$par

}
else if(distribution == "lognormal"){
  par_1 <- optim(par=c(2,0.4,0.8),Loglik,method='L-BFGS-B',
                 lower=c(1,0.1,0),upper=c(5,1,1))$par
}

#model parameter and pi

a = par_1[1]
b = par_1[2]
p = par_1[3]

a.vec[run] = a
b.vec[run] = b
q.vec[run] = 1-p



#mean, median, quartiles
if(distribution == "gamma"){
  ##########################gamma
  est <- c(
    c(a,b,1-p) ,  # estimates for parameters alpha and beta, and
       pi
    a/b,          # mean
    round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
       estimated quantiles
  # -round(loglik(c(a,b)),2),

} else if(distribution == "weibull"){
  ##########################weibull
  est <- c(
    c(a,b,1-p) ,   # estimates for parameters k and lambda, and
       pi
    b*gamma(1+1/a) , # mean
```

```
    round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
      ),2))  # estimated quantiles
  #-round(loglik(c(a,b)),2),

} else if(distribution == "lognormal"){
  ########################lognormal
  est <- c(
    c(a,b,1-p),  # estimates for parameters mu and sigma, and pi
    exp(a+b^2/2), # mean
    round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
      estimated quantiles
  #-round(loglik(c(a,b)),2)

}

#mean, median, quartiles, bias

ave.vec[run] = est[4]
Q1.vec[run] = est[5]
Q2.vec[run] = est[6]
Q3.vec[run] = est[7]
Q4.vec[run] = est[8]
Q5.vec[run] = est[9]
Q6.vec[run] = est[10]


bias_par1_Deng[run] <- est[1] - aa
bias_par2_Deng[run] <- est[2] - bb
bias_p_Deng[run] <- est[3] - (1 - pp)



if(distribution == "gamma"){
  par_2 <- optim(par=c(4,0.5,0.8),loglik,method='L-BFGS-B',
                 lower=c(1.1,0.1,0),upper=c(10,2,1))$par
}
else if(distribution == "weibull"){
  par_2 <- optim(par=c(2,10,0.8),loglik,method='L-BFGS-B',
                 lower=c(1.1,2,0),upper=c(5,15,1))$par
}
else if(distribution == "lognormal"){
  par_2 <- optim(par=c(2,0.4,0.8),loglik,method='L-BFGS-B',
                 lower=c(1,0.1,0),upper=c(5,1,1))$par
}

#model parameter and pi

a = par_2[1]
b = par_2[2]
p = par_2[3]

a.vec.qin[run] = a
b.vec.qin[run] = b
q.vec.qin[run] = 1-p


#mean, median, quartiles
if(distribution == "gamma"){
```

```
    ##########################gamma
    est <- c(
      c(a,b,1-p),  # estimates for parameters alpha and beta, and
          pi
      a/b,         # mean
      round(qgamma(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b),2))  #
          estimated quantiles
    # -round(loglik(c(a,b)),2),

  } else if(distribution == "weibull"){
    ##########################weibull
    est <- c(
      c(a,b,1-p),   # estimates for parameters k and lambda, and pi
      b*gamma(1+1/a), # mean
      round(qweibull(c(0.25,0.5,0.75,0.9,0.95,0.99),shape=a,scale=b
          ),2))  # estimated quantiles
    #-round(loglik(c(a,b)),2),

  } else if(distribution == "lognormal"){
    ########################lognormal
    est <- c(
      c(a,b,1-p),  # estimates for parameters mu and sigma, and pi
      exp(a+b^2/2), # mean
      round(exp(qnorm(c(0.25,0.5,0.75,0.9,0.95,0.99),a,b)),2))  #
          estimated quantiles
    #-round(loglik(c(a,b)),2)

  }


  #mean, median, quartiles, bias

  ave.vec.qin[run] = est[4]
  Q1.vec.qin[run] = est[5]
  Q2.vec.qin[run] = est[6]
  Q3.vec.qin[run] = est[7]
  Q4.vec.qin[run] = est[8]
  Q5.vec.qin[run] = est[9]
  Q6.vec.qin[run] = est[10]


  bias_par1_Qin[run] <- est[1] - aa
  bias_par2_Qin[run] <- est[2] - bb
  bias_p_Qin[run] <- est[3] - (1 - pp)


}

#MSE and coverage
MSE_par1_Deng <- sum(bias_par1_Deng^2)/runs
MSE_par2_Deng <- sum(bias_par2_Deng^2)/runs
MSE_p_Deng <- sum(bias_p_Deng^2)/runs

MSE_par1_Qin <- sum(bias_par1_Qin^2)/runs
MSE_par2_Qin <- sum(bias_par2_Qin^2)/runs
MSE_p_Qin <- sum(bias_p_Qin^2)/runs
```

```
res <- rbind(c(round(mean(a.vec),2), round(t.test(a.vec)$conf.int
    [1], 2), round(t.test(a.vec)$conf.int[2], 2)),
             c(round(mean(b.vec),2), round(t.test(b.vec)$conf.int
                [1], 2), round(t.test(b.vec)$conf.int[2], 2)),
             c(round(mean(q.vec),2), round(t.test(q.vec)$conf.int
                [1], 2), round(t.test(q.vec)$conf.int[2], 2)),
             c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
             c(round(mean(Q1.vec),2), round(t.test(Q1.vec)$conf.int
                [1], 2), round(t.test(Q1.vec)$conf.int[2], 2)),
             c(round(mean(Q2.vec),2), round(t.test(Q2.vec)$conf.int
                [1], 2), round(t.test(Q2.vec)$conf.int[2], 2)),
             c(round(mean(Q3.vec),2), round(t.test(Q3.vec)$conf.int
                [1], 2), round(t.test(Q3.vec)$conf.int[2], 2)),
             c(round(mean(Q4.vec),2), round(t.test(Q4.vec)$conf.int
                [1], 2), round(t.test(Q4.vec)$conf.int[2], 2)),
             c(round(mean(Q5.vec),2), round(t.test(Q5.vec)$conf.int
                [1], 2), round(t.test(Q5.vec)$conf.int[2], 2)),
             c(round(mean(Q6.vec),2), round(t.test(Q6.vec)$conf.int
                [1], 2), round(t.test(Q6.vec)$conf.int[2], 2)),
             c(round(MSE_par1_Deng, 2), NULL, NULL),
             c(round(MSE_par2_Deng, 2), NULL, NULL),
             c(round(MSE_p_Deng, 2), NULL, NULL),

             c(round(mean(a.vec.qin),2), round(t.test(a.vec.qin)$
                conf.int[1], 2), round(t.test(a.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(b.vec.qin),2), round(t.test(b.vec.qin)$
                conf.int[1], 2), round(t.test(b.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(q.vec.qin),2), round(t.test(q.vec.qin)$
                conf.int[1], 2), round(t.test(q.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(ave.vec),2), round(t.test(ave.vec)$conf.
                int[1], 2), round(t.test(ave.vec)$conf.int[2], 2)),
             c(round(mean(Q1.vec.qin),2), round(t.test(Q1.vec.qin)$
                conf.int[1], 2), round(t.test(Q1.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(Q2.vec.qin),2), round(t.test(Q2.vec.qin)$
                conf.int[1], 2), round(t.test(Q2.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(Q3.vec.qin),2), round(t.test(Q3.vec.qin)$
                conf.int[1], 2), round(t.test(Q3.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(Q4.vec.qin),2), round(t.test(Q4.vec.qin)$
                conf.int[1], 2), round(t.test(Q4.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(Q5.vec.qin),2), round(t.test(Q5.vec.qin)$
                conf.int[1], 2), round(t.test(Q5.vec.qin)$conf.int
                [2], 2)),
             c(round(mean(Q6.vec.qin),2), round(t.test(Q6.vec.qin)$
                conf.int[1], 2), round(t.test(Q6.vec.qin)$conf.int
                [2], 2)),
             c(round(MSE_par1_Qin, 2), NULL, NULL),
             c(round(MSE_par2_Qin, 2), NULL, NULL),
             c(round(MSE_p_Qin, 2), NULL, NULL))
```

```
    parameter <- c('alpha','beta', 'pi', 'mean', '.25Q', 'median', '.75
       Q', '.90Q', '.95Q', '.99Q', 'MSE alpha', 'MSE beta', 'MSE p',
                  'alpha','beta', 'pi', 'mean', '.25Q', 'median', '.75
                     Q', '.90Q', '.95Q', '.99Q', 'MSE alpha', 'MSE
                     beta', 'MSE p')
    method <- c(rep('Deng', nrow(res) / 2), rep('Qin', nrow(res) / 2))

    colnames(res) <- c("estimate", "L_95%CI", "U_95%CI")

    res <- data.frame(parameter, method, res)

    result = rbind(result, res)
    print(cat("sample size:", m, "pi:", 1-pp))
    print(res)
  }
  return(result)
}
```