FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Petr Míchal

Gradual change model

Department of Probability and Mathematical Statistics

Supervisor of the master thesis: doc. RNDr. Zdeněk Hlávka, Ph.D.
Study programme: Mathematics
Study branch: Probability, Mathematical Statistics and Econometrics

Prague 2020

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

I would like to express my gratitude to my supervisor doc. RNDr. Zdeněk Hlávka, Ph.D. for his support, helpful comments and his guidance through all the stages of writing. I would also like to thank Mgr. Martin Otava, Ph.D. for his reviews and comments.

Title: Gradual change model

Author: Petr Míchal

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Zdeněk Hlávka, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The thesis aims at change-point estimation in gradual change models. Methods available in literature are reviewed and modified for point-of-stabilisation (*PoSt*) context, present e.g. in drug continuous manufacturing. We describe in detail the estimation in the linear *PoSt* model and we extend the methods to quadratic and E_{max} model. We describe construction of confidence intervals for the change-point, discuss their interpretation and show how they can be used in practice. We also address the situation when the assumption of homoscedasticity is not fulfilled. Next, we run simulations to calculate the coverage of confidence intervals for the change-point in discussed models using asymptotic results and bootstrap with different parameter combinations. We also inspect the simulated distribution of derived estimators with finite sample. In the last chapter, we discuss the situation when the model for the data is incorrectly specified and we calculate the coverage of confidence intervals using simulations.

Keywords: change-point analysis, gradual change, E_{max} model, point-of-stabilization

Contents

Introduction

In change-point analysis, there are two main tasks, testing a presence of a change-point in data and estimating the change-point and other parameters of assumed model, while the change can be abrupt or gradual. The thesis aims at estimation in gradual change model. In such models, the change appears gradually, e.g. the mean value of the outcome changes from constant to linear after the change-point. Such behaviour appears often in real world processes, e.g. in continuous manufacturing, where quality of the products is not the same because of the start-up period of the production line. After some time, the process stabilises and the expected quality of the product does not show any trend. In such scenario, the trend is present at the beginning up to the change-point and the process stabilises after the change-point, i.e. the mean value of the outcome becomes constant.

It is important to estimate the point-of-stabilisation (the change-point) in order to guarantee the same quality of the products and to minimise waste of material during the start-up phase. In this situation, we want to estimate the change-point and the other parameters of the model and construct the confidence interval for the change-point, either using the asymptotic results or using bootstrap approximation.

We modify results from Hušková [1998], Hlávka and Hušková [2017] and Jarušková [2001] to fit into the *PoSt* context, namely we change the time ordering in linear model from Hušková [1998] and Hlávka and Hušková [2017] and we assume general variance of the random errors in quadratic model from Jarušková [2001]. Further, we introduce a nonpolynomial model with a change-point, namely the E_{max} model. In comparison to the quadratic model, the E_{max} model keeps monotonicity which is a common assumption in various scientific applications. In a quadratic model it sometimes happen that the trend changes its monotonicity near the change-point. We show how to construct confidence intervals for the change-point using asymptotic results or bootstrap and we discuss how to interpret and use them in practice to verify the stability of the process. Also, we simulate the coverage of confidence intervals based on the asymptotic results and bootstrap for different locations of the change-point and sample sizes and we compare both methods. We also explore what happens, when the model is incorrectly specified.

In Chapter 1, we describe methods for testing the presence of the change-point in various models and methods for estimation of the change-point and other parameters available in literature.

Next, we aim at estimation in polynomial change models in Chapter 2. We describe the estimation using least squares method in gradual change models with arbitrary polynomial trend and we state general formulae for the estimators. For the linear model, the asymptotic results were derived in Hušková [1998], the results for the quadratic model in Jarušková [2001]. We introduce the E_{max} model (which is used in dose-response studies, see e.g.MacDougall [2006]), by including a change-point into the model and we derive estimators of the unknown parameters in this model.

In Chapter 3, we introduce the point-of-stabilisation model which can be used e.g. in drug continuous manufacturing, where it captures the product output quality containing a trend during a start-up period of the production line and after the stabilisation. In this context, the change-point represents the time the production line stabilises, so called point-of-stabilisation (*PoSt*). We briefly discuss testing in *PoSt* model and the differences against testing in linear gradual change model discussed in previous chapter. Next, we aim at estimation of unknown parameters in the model, we modify the formulae from previous sections to take into account time ordering in *PoSt* context and we state the asymptotic distribution of modified estimators. We construct confidence intervals for the change-point and discuss their connection to testing the stability of the production process in practice. Next, we run simulations to verify asymptotic results, we compare the asymptotic distribution of estimators with the simulated distribution with finite sample sizes. We also calculate the coverage of confidence intervals for the change-point for more parameter combinations using both the asymptotic distribution and bootstrap approximation.

Next, in Chapter 4 we discuss the case when homoscedasticity (which is assumed in previous models) is not fulfilled and we show how to modify the estimators to take heteroscedasticity of the random errors into account by assuming multiple measurements at each time i to be able to estimate the variance for each time i. We show the method on the linear *PoSt* model, but it applies analogously also to other models.

In Chapter 5, we generalise the linear *PoSt* model by assuming more complicated trend than linear before the change-point. First, we discuss the quadratic *PoSt* model, we run simulations to compare the asymptotic and the simulated distribution of the estimators. We show how to construct confidence intervals and we calculate their coverage for both methods. Then, we focus on the E_{max} *PoSt* model introduced in Section 2.3. For this model, we show the simulated distribution of the estimators since the asymptotic results for this model with change-point are not available and we calculate the coverage of confidence intervals constructed using bootstrap.

In Chapter 6, we explore what happens, when the model for the data is incorrectly specified and the variance structure of the errors (heteroscedasticity or homoscedasticity) is assumed incorrectly, which can often happen in reality and it should be explored. We calculate the coverage of confidence intervals for the change-point for more locations of the change-point. In the first scenario, the assumed model is more complex than the true model. In the second scenario, the situation is inverse, the true model is more complicated than the assumed model.

1. Gradual change model

Change-point analysis is a part of statistical analysis examining a situation when the underlying probability distribution of data changes in time. The change can be abrupt (e.g. jump in mean value) or gradual, which will be our case. Gradual change model represents a situation when a trend in data gradually changes or appears at unknown change-point. In the usual setup, the expectation is assumed to be constant up to an unknown change-point κ. After κ, a monotonic trend starts to appear. For example, the expected value can be constant up to κ and it starts following a linear trend after κ, as in Figure 1.1.

Let us assume that observations $Y_1, \ldots, Y_n$ follow polynomial change-point model with unknown change-point κ

$$Y_i = \beta_0 + \beta_1 \left(\frac{i-\kappa}{n}\right)^+ + \beta_2 \left(\left(\frac{i-\kappa}{n}\right)^+\right)^2 + \cdots + \beta_d \left(\left(\frac{i-\kappa}{n}\right)^+\right)^d + e_i, \quad (1.1)$$

where $d \in \mathbb{N}$, c^+ denotes positive part of c, i.e. $c^+ = \max\{0, c\}$, $i = 1, \ldots, n$, Random errors $e_1, \ldots, e_n$ are iid and satisfy $\mathsf{E}\, e_i = 0$, $\operatorname{var} e_i = \sigma^2 > 0$ and $\mathsf{E}\, |e_i|^{2+\Delta} < \infty$ for some $\Delta > 0$. The parameter d represents the degree of polynomial trend after change-point κ. For $i \leq \kappa$ we have $Y_i = \beta_0 + e_i$.

One of the main tasks concerning model (1.1) is finding the asymptotic distribution of estimators of the unknown parameters of the model. The second task is testing a presence of the change-point.

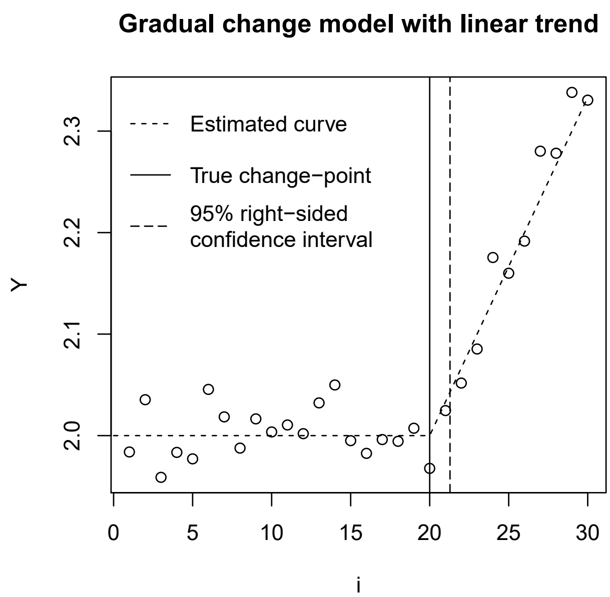

Figure 1.1: Gradual change model with linear trend and with right-sided asymptotic 95% confidence interval for change-point κ given by (2.9).

1.1 Testing

Testing the presence of change-point can be viewed as testing the null hypothesis $H_0 : \kappa = n$ (there is no change-point and constant model holds) against the alternative $H_1 : \kappa < n$. Jarušková [1998b] developed a testing procedure in gradual change model (1.1) with $d = 1$. Testing in a more general model

$$Y_i = \mu + \delta \left(\left(\frac{i-\kappa}{n} \right)^+ \right)^\alpha + e_i$$

for some known $\alpha > 0$ was discussed in Hušková and Steinebach [2000]. Unlike d in model (1.1), the parameter α was assumed to be continuous. For $\alpha = 0$, the change is abrupt and for $\alpha = 1$, the model is equivalent to (1.1) with $d = 1$.

In Rusá [2015], testing a presence of change-point in panel data setup was examined and test statistics for testing the change in trend was developed. We can imagine panel data as a situation when N subjects are followed over period of time T. The author assumed the data to be in form $X_{it},\ i = 1, \ldots, N,\ t = 1, \ldots, T,$ where the observation X_{it} was measured on i-th subject at time t. The author developed tests for testing the presence of the change-point t_0 in such data when assuming a linear trend in time which changes after the change-point, i.e.

$$X_{it} = \mu_i + \beta_i t + \delta_i (t - t_0)^+ + e_{it}, \quad i = 1, \ldots, N,\ t = 1, \ldots, T,\ 1 < t_0 < T,$$

where μ_i, γ_i are unknown parameters and e_i are random errors.

1.2 Estimation

Parameter estimation together with determining the asymptotic distribution in model (1.1) for the case $d = 1$ (linear trend) was discussed in Hušková [1998]. The same results were derived in Jarušková [1998a] as a special case of a more general model. Hušková [1999] derived the asymptotic distribution of the least-squares estimators for more general case $d = 1$ and $\left[\left((i-\kappa)/n \right)^+ \right]^\alpha$ for known $\alpha > 0$ instead of $\left((i-\kappa)/n \right)^+$, for some known $\alpha > 0$.

Estimation with quadratic trend ($d = 2$) was discussed in Jarušková [1998a] and in Jarušková [2001]. Jarušková [1998a] worked with model

$$Y_i = \alpha_0 + \alpha_1 \left(\frac{i}{n} \right) + \cdots + \alpha_p \left(\frac{i}{n} \right)^p + \beta \left(\left(\frac{i-\kappa}{n} \right)^+ \right)^q + e_i, \quad i = 1, \ldots, n,$$

for some known $p = 0, 1, \ldots, q > 1$ and random error e_i as in (1.1). This model represents a situation when the change affects only the highest degree of polynomial trend and the other coefficients are nuisance parameters. The author derived estimators for this case together with their asymptotic distribution. Linear trend discussed in Hušková [1998] is a special case of this model.

In Jarušková [2001], the model captured the change in both the linear and quadratic term. On the other hand, the author assumed the parameters describing the expected value before the change-point to be known and without loss of generality set to zero, leading to

$$Y_i = \beta \left(\frac{i-\kappa}{n} \right)^+ + \gamma \left(\left(\frac{i-\kappa}{n} \right)^+ \right)^2 + e_i.$$

The asymptotic distribution of unknown parameters κ, β, γ was derived. Also, a small simulation study concerning the limit distribution was done.

In Döring [2015], the model represented a situation with asymmetric regression function with change at unknown change-point θ. Both parts before and after θ could have different degree of smoothness. Specifically, the regression function had form

$$\begin{aligned} f_{\theta,p,q,\boldsymbol{a}}(x) &= g_0(x, \boldsymbol{a}) \cdot \mathbb{1}_{[0,1]}(x) \\ &\quad + g_1(x, \boldsymbol{a}) \cdot (\theta - x)^p \, \mathbb{1}_{[0,\theta)}(x) + g_2(x, \boldsymbol{a}) \cdot (x - \theta)^q \, \mathbb{1}_{(\theta,1]}(x), \end{aligned}$$

where $\theta \in [0, 1]$ denotes change-point, $p, q \in [0, \infty)$ are degrees of smoothness and $\boldsymbol{a} \in \mathbb{R}^d$ represents a vector of nuisance parameters. Further, functions g_0, g_1, $g_2 : R^{d+1} \to \mathbb{R}$ were assumed to be two times continuously differentiable. The behaviour of least squares estimators of $(\theta, p, q, \boldsymbol{a})$ was studied, based on observations (X_i, Y_i), $i = 1, \ldots, n$, where $Y_i = f_{\theta,p,q,\boldsymbol{a}}(X_i) + e_i$ for each i. Random errors e_i were assumed to be iid with $\mathsf{E}\,(e_i|X) = 0$ a.s. and suitably integrable. Consistency of estimators and their limit behaviour was then studied and it turned out it depends on $b = \min(p, q)$. For $b \geq \frac{1}{2}$ the derived estimators were asymptotically normal with higher rate of convergence of the change-point estimator in case $b = \frac{1}{2}$. For $b < \frac{1}{2}$, the asymptotic distribution can be represented as a unique maximiser of a fractional Brownian motion with drift.

Model (1.1) with $d = 1$ is a special case of this situation with $g_0 = \beta_0$, $g_1 = 0$, $g_2 = \beta_1$, $X_i = i/n$, $\theta = \kappa/n$ and $q = 1$.

2. Estimation in gradual change model

Model of gradual change can be used in various ways, e.g. in industry and in meteorological measurements. In this chapter, we discuss estimation in polynomial change model. In linear gradual change model, we present the asymptotic results derived in Hušková [1998], we construct confidence intervals for the change-point and shortly discuss their interpretation. Next, we move to quadratic model and we introduce the E_{max} model.

For simplicity, let us define

$$x_{i,k} = \left(\frac{i-k}{n}\right)^{+}, \qquad i = 1, \dots, n;\ k \in (1, n)$$

$$\overline{x}_{\cdot k} = \frac{1}{n}\sum_{i=1}^{n} x_{i,k}, \qquad k \in (1, n)\,.$$

Model (1.1) has unknown parameters $\boldsymbol{\beta} = (\beta_0, \dots, \beta_d)^\top, \sigma^2$ and κ. Parameters $\boldsymbol{\beta}, \kappa$ can be estimated by least squares method. The estimators are given as a solution of minimization problem

$$\min_{\substack{\beta_0,\dots,\beta_d \in \mathbb{R} \\ k \in (1,n)}} \quad \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 x_{i,k} - \dots - \beta_d x_{i,k}^d\right)^2 .$$

Denoting

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \qquad \mathbb{X}_{\cdot k} = \begin{pmatrix} 1 & x_{1,k} & x_{1,k}^2 & \dots & x_{1,k}^d \\ & \vdots & & & \\ 1 & x_{n,k} & x_{n,k}^2 & \dots & x_{n,k}^d \end{pmatrix},$$

we can rewrite our minimization task as

$$\min_{\substack{\beta_0,\dots,\beta_d \in \mathbb{R} \\ k \in (1,n)}} \quad \|\boldsymbol{Y} - \mathbb{X}_{\cdot k}\boldsymbol{\beta}\| = \min_{\substack{\beta_0,\dots,\beta_d \in \mathbb{R} \\ k \in (1,n)}} \quad \left(\boldsymbol{Y} - \mathbb{X}_{\cdot k}\boldsymbol{\beta}\right)^\top \left(\boldsymbol{Y} - \mathbb{X}_{\cdot k}\boldsymbol{\beta}\right). \tag{2.1}$$

Direct calculations give specific forms of the estimators of $\boldsymbol{\beta}, \kappa$. We have

$$\begin{aligned} \widehat{\kappa} &= \underset{k \in (1,n)}{\arg\min}\ \boldsymbol{Y}^\top \left(\mathbb{I} - \mathbb{X}_{\cdot k}\left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top\right) \boldsymbol{Y} \\ &= \underset{k \in (1,n)}{\arg\min}\ \boldsymbol{Y}^\top \boldsymbol{Y} - \boldsymbol{Y}^\top \mathbb{X}_{\cdot k}\left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top \boldsymbol{Y} \\ &= \underset{k \in (1,n)}{\arg\max}\ \boldsymbol{Y}^\top \mathbb{X}_{\cdot k}\left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top \boldsymbol{Y}\,. \end{aligned} \tag{2.2}$$

Remark. Estimation of the change-point can be equivalently done using coefficient of determination. For given $k \in (1, n)$, assume a linear model with response $\boldsymbol{Y}$ and model matrix $\mathbb{X}_{\cdot k}$. Denote R_k^2 the coefficient of determination of the model and $\widehat{\boldsymbol{Y}} = \mathbb{X}_{\cdot k}\left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top \boldsymbol{Y}$ fitted values. Then

$$R_k^2 = 1 - \frac{\sum_{i=1}^n \left(Y_i - \widehat{Y}_i\right)^2}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2} = 1 - \frac{\boldsymbol{Y}^\top \boldsymbol{Y} - \boldsymbol{Y}^\top \mathbb{X}_{\cdot k} \left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top \boldsymbol{Y}}{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2}.$$

The expression $\boldsymbol{Y}^\top \boldsymbol{Y} - \boldsymbol{Y}^\top \mathbb{X}_{\cdot k} \left(\mathbb{X}_{\cdot k}^\top \mathbb{X}_{\cdot k}\right)^{-1} \mathbb{X}_{\cdot k}^\top \boldsymbol{Y}$ is minimised in (2.2). Using the equation above, we rewrite the argument of minimisation and we obtain an equivalent formula to estimate the change-point:

$$\widehat{\kappa} = \underset{k \in (1,n)}{\arg\min} \ \left(1 - R_k^2\right) \sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2 = \underset{k \in (1,n)}{\arg\max} \ R_k^2 \,. \tag{2.3}$$

Vector of parameters $\boldsymbol{\beta}$ can be estimated by

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\cdot \widehat{\kappa}}^\top \mathbb{X}_{\cdot \widehat{\kappa}}\right)^{-1} \mathbb{X}_{\cdot \widehat{\kappa}}^\top \boldsymbol{Y} \,. \tag{2.4}$$

The parameter σ^2 can be estimated by

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{i,\widehat{\kappa}} - \cdots - \widehat{\beta}_d x_{i,\widehat{\kappa}}^d\right)^2 \,. \tag{2.5}$$

The formulae hold also for a situation with general matrix $\mathbb{X}$ depending on k, i.e.

$$\mathbb{X}_k = \begin{pmatrix} 1 & \boldsymbol{x}_1(k)^\top \\ & \vdots \\ 1 & \boldsymbol{x}_n(k)^\top \end{pmatrix},$$

for vectors $\boldsymbol{x}_i(k) \in \mathbb{R}^d, \ i = 1, \ldots, n$ depending on k. This case will be discussed in Section 2.3.

2.1 Linear trend

Assume the data $Y_1, \ldots, Y_n$ satisfy for each $i = 1, \ldots, n$

$$Y_i = \beta_0 + \beta_1 \, x_{i,\kappa} + e_i = \beta_0 + \beta_1 \left(\frac{i - \kappa}{n}\right)^+ + e_i, \tag{2.6}$$

where random errors e_i are as in model (1.1) and $\kappa \in \{1, \ldots, n\}$. Estimation and the asymptotic distribution of estimators in this model was discussed in Hušková [1998]. Similarly as in Hlávka and Hušková [2017], we estimate κ on a continuous scale by

$$\widehat{\kappa} = \underset{k \in (1,n)}{\arg\max} \ \frac{\left(\sum_{i=1}^n Y_i \left(x_{i,k} - \overline{x}_{\cdot k}\right)\right)^2}{\sum_{i=1}^n \left(x_{i,k} - \overline{x}_{\cdot k}\right)^2}, \tag{2.7}$$

which is equivalent to (2.2) for $d = 1$.

Estimators of β_0, β_1 are given by (2.4). In the assumed model, they can also be expressed as

$$\widehat{\beta}_0 = \overline{Y}_n - \widehat{\beta}_1 \, \overline{x}_{\cdot\widehat{\kappa}}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n Y_i \left(x_{i,\widehat{\kappa}} - \overline{x}_{\cdot\widehat{\kappa}} \right)}{\sum_{i=1}^n \left(x_{i,\widehat{\kappa}} - \overline{x}_{\cdot\widehat{\kappa}} \right)^2}. \tag{2.8}$$

The estimator $\widehat{\kappa}$ can be equivalently calculated using a coefficient of determination R^2 as in remark in previous section.

The parameter σ^2 can be estimated by

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{i,\widehat{\kappa}} \right)^2 .$$

Hušková [1998] derived the asymptotic distribution of estimators $\widehat{\kappa}$, $\widehat{\beta}_0$ and $\widehat{\beta}_1$ in this model.

Theorem 1. *Assume $Y_1, \dots, Y_n$ are independent and satisfy model* (2.6)*. Let, as $n \to \infty$,*

$$\beta_1 = O(1), \qquad \frac{\beta_1^2 n}{(\log \log n)} \longrightarrow \infty$$

and

$$\kappa = [n\theta]$$

for some $\theta \in (0, 1)$.

Then, as $n \to \infty$,

$$\frac{\beta_1}{\sigma} \frac{\widehat{\kappa} - \kappa}{\sqrt{n}} \sqrt{\frac{\theta\,(1-\theta)}{1+3\theta}} \xrightarrow{D} \mathrm{N}(0, 1).$$

Proof. Hušková [1998, Theorem A].

□

The asymptotic distribution of the estimator $\widehat{\kappa}$ can be used to construct asymptotic confidence intervals for the change-point κ.

Often, one-sided confidence intervals are desired, because of their interpretation and connection to testing the stability of the production process, which will be further discussed in Chapter 3. From Theorem 1, we obtain the right-sided confidence interval

$$(-\infty, c_U) = \left(-\infty, \; \widehat{\kappa} + u_{1-\alpha} \frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta}_1} \sqrt{\frac{1+3\widehat{\theta}}{\widehat{\theta}(1-\widehat{\theta})}} \right), \tag{2.9}$$

where u_α denotes the α-quantile of $\mathrm{N}(0, 1)$ and $\widehat{\theta} = \widehat{\kappa}/n$. The time c_U can be interpreted as the time after which the mean value of Y_i significantly differs from β_0, see Figure 1.1. From duality of confidence intervals and hypothesis testing, this confidence interval is connected to testing the null hypothesis H_0 against the alternative H_1, where

$$H_0 : \kappa \geq \kappa_0$$

$$H_1 : \kappa < \kappa_0$$

for some constant κ_0. We reject H_0 if $\kappa_0 \notin (-\infty, c_U)$.

It holds $\mathsf{E}\, Y_i = \beta_0$ for $i = 1, \ldots, \kappa$ and $\mathsf{E}\, Y_i = \beta_0 + \beta_1 x_{i,\kappa}$ for $i = \kappa, \ldots, n$. Therefore, $\kappa > \kappa_0$ means the trend does not influence Y_{κ_0} since the change-point occurs after κ_0, see Figure 1.1. We can equivalently formulate hypotheses above as

$$H_0 : \mathsf{E}\, Y_{\kappa_0} = \beta_0$$
$$H_1 : \mathsf{E}\, Y_{\kappa_0} \neq \beta_0.$$

Similarly, left-sided confidence interval

$$(c_L, \infty) = \left(\widehat{\kappa} - u_{1-\alpha} \frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta_1}} \sqrt{\frac{1 + 3\widehat{\theta}}{\widehat{\theta}(1 - \widehat{\theta})}},\ \infty \right),$$

is connected to testing

$$H_0 : \kappa \leq \kappa_0$$
$$H_1 : \kappa > \kappa_0$$

for some κ_0 and rejecting H_0 if $\kappa_0 \notin (c_L, \infty)$. The interpretation of the confidence intervals, the connection to testing and their use in practice will be further discussed in *point-of-stabilisation* context in Chapter 3.

2.2 Quadratic trend with reversed time

In reality, the data usually follow more complicated trend than linear. We will now focus on model with quadratic trend. Moreover, the model will be formulated with „reversed“ time ordering similarly as in Jarušková [2001] which will be further used in Chapter 3 concerning *PoSt* model. Unlike in previous section, here the trend is present up to the change-point and after the change-point the data do not show any trend. For clarity, we will denote the change-point in the „reversed“ context by ψ instead of κ and the data by Z_i instead of Y_i.

Assume we have data $Z_1, \ldots, Z_n$ from model

$$Z_i = \beta_0 + \beta_1 \left(\frac{\psi - i}{n} \right)^+ + \beta_2 \left(\left(\frac{\psi - i}{n} \right)^+ \right)^2 + e_i, \tag{2.10}$$

where random errors $e_1, \ldots, e_n$ satisfy $\mathsf{E}\, e_i = 0$, $\operatorname{var} e_i = \sigma^2 > 0$ and we have $\psi \in \{1, \ldots, n\}$. Unknown parameters are β_0, β_1, β_2, ψ and σ^2. This model represents the situation when data follow a quadratic trend up to an unknown change-point ψ and become stable after ψ. In our model, both the linear and the quadratic term are present up to ψ, unlike in Jarušková [1998a], where the change occurred only at the quadratic term.

We distinguish two situations depending on whether β_0 is known or not, since the asymptotic distributions differ.

2.2.1 Known β_0

When β_0 is known, we can assume without loss of generality that $\beta_0 = 0$, otherwise we could work with $\widetilde{Z}_i = Z_i - \beta_0$, $i = 1, \ldots, n$. The model (2.10) simplifies to

$$Z_i = \beta_1 \left(\frac{\psi - i}{n}\right)^+ + \beta_2 \left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + e_i, \tag{2.11}$$

where random errors e_i are as in (2.10). This model was studied in Jarušková [2001] with known $\sigma^2 = 1$. Denote $x_{p,i}^s = \left(\left(\frac{\psi-i}{n}\right)^+\right)^s$ for $s = 1, 2$ and

$$\mathbb{X}_{p\cdot} = \begin{pmatrix} x_{p,1} & x_{p,1}^2 \\ x_{p,2} & x_{p,2}^2 \\ \vdots & \vdots \\ x_{p,n} & x_{p,n}^2 \end{pmatrix}.$$

Point estimates can be derived similarly as in previous chapters. We have

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max}\ \boldsymbol{Z}^\top \mathbb{X}_{p\cdot} \left(\mathbb{X}_{p\cdot}^\top \mathbb{X}_{p\cdot}\right)^{-1} \mathbb{X}_{p\cdot}^\top \boldsymbol{Z}$$

or, while denoting R_p^2 the coefficient of determination of the linear model with response $\boldsymbol{Y}$ and model matrix $\mathbb{X}_{p\cdot}$, as

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max}\ R_p^2 . \tag{2.12}$$

The vector of parameters $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top$ can be estimated similarly as before by

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\widehat{\psi}\cdot}^\top \mathbb{X}_{\widehat{\psi}\cdot}\right)^{-1} \mathbb{X}_{\widehat{\psi}\cdot}^\top \boldsymbol{Z}.$$

The asymptotic distribution of the estimators differs depending on β_1. It is normal for the case $\beta_1 \neq 0$. If $\beta_1 = 0$ we obtain non-normal asymptotic distribution, see Jarušková [2001]. Moreover, we have to deal with unknown variance σ^2.

Let $\theta_\psi = \psi/n \in [\delta, 1-\delta]$ for a known constant $\delta \in (0, 1/2)$ and $\widehat{\theta}_\psi = \widehat{\psi}/n$.

Theorem 2. *Suppose model* (2.11) *holds and* $\beta_1 \neq 0$. *Then*

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi, \widehat{\beta}_1 - \beta_1, \widehat{\beta}_2 - \beta_2\right)^\top$$

has asymptotically a zero-mean normal distribution with a variance-covariance matrix $\mathbb{G}$, *where*

$$\mathbb{G} = \begin{pmatrix} \frac{9\sigma^2}{\beta_1^2\theta_\psi} & -\frac{\left(36\beta_1 - 18\beta_2\theta_\psi\right)\sigma^2}{\beta_1^2\theta_\psi^2} & \frac{30\beta_1\sigma^2}{\beta_1^2\theta_\psi^3} \\ -\frac{\left(36\beta_1 - 18\beta_2\theta_\psi\right)\sigma^2}{\beta_1^2\theta_\psi^2} & \frac{\left(36\beta_2^2\theta_\psi^2 + 144\beta_1\beta_2\theta_\psi + 192\beta_1^2\right)\sigma^2}{\beta_1^2\theta_\psi^3} & -\frac{\left(180\beta_1 + 60\beta_2\theta_\psi\right)\beta_1\theta_\psi\sigma^2}{\beta_1^2\theta_\psi^5} \\ \frac{30\beta_1\sigma^2}{\beta_1^2\theta_\psi^2} & -\frac{\left(180\beta_1 + 60\beta_2\theta_\psi\right)\beta_1\theta_\psi\sigma^2}{\beta_1^2\theta_\psi^5} & \frac{180\sigma^2}{\theta_\psi^5} \end{pmatrix}.$$

Proof. We will use Theorem A from Jarušková [2001] but we have to take into account more general variance of random errors e_i than $\sigma^2 = 1$.

Define $Z_i^* = \frac{Z_i}{\sigma}$. Using definition of model (2.11), we have

$$Z_i^* = \frac{\beta_1}{\sigma}\left(\frac{\psi - i}{n}\right)^+ + \frac{\beta_2}{\sigma}\left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + \frac{e_i}{\sigma}$$

$$= \beta_1^*\left(\frac{\psi - i}{n}\right)^+ + \beta_2^*\left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + e_i^*$$

denoting $\beta_i^* = \beta_i/\sigma$ and $e_i^* = e_i/\sigma$. We have $\operatorname{var} e_i^* = 1$ and the matrix $\mathbb{X}_{p.}$ does not change neither does the change-point ψ. Also, the data $Z_1^*, \ldots, Z_n^*$ satisfy the model used in Jarušková [2001], which is the same as our model (2.11) but having random errors with variance equal to 1.

One can estimate $\boldsymbol{\beta}^*$ and ψ from $Z_1^*, \ldots, Z_n^*$ as usually. We have

$$\widehat{\psi} = \underset{p\,\in\,(1,n)}{\arg\max}\ \boldsymbol{Z}^{*\top}\mathbb{X}_{p.}\left(\mathbb{X}_{p.}^\top\mathbb{X}_{p.}\right)^{-1}\mathbb{X}_{p.}^\top\boldsymbol{Z}^*$$

$$= \underset{p\,\in\,(1,n)}{\arg\max}\ \boldsymbol{Z}^{\top}\mathbb{X}_{p.}\left(\mathbb{X}_{p.}^\top\mathbb{X}_{p.}\right)^{-1}\mathbb{X}_{p.}^\top\boldsymbol{Z}$$

and

$$\widehat{\boldsymbol{\beta}}^* = \left(\mathbb{X}_{\widehat{\psi}.}^\top\mathbb{X}_{\widehat{\psi}.}\right)^{-1}\mathbb{X}_{\widehat{\psi}.}^\top\boldsymbol{Z}^* = \left(\mathbb{X}_{\widehat{\psi}.}^\top\mathbb{X}_{\widehat{\psi}.}\right)^{-1}\mathbb{X}_{\widehat{\psi}.}^\top\boldsymbol{Z}/\sigma = \frac{\widehat{\boldsymbol{\beta}}}{\sigma}.$$

Using Theorem A from Jarušková [2001] we obtain

$$\sqrt{n}\begin{pmatrix}\widehat{\theta}_\psi - \theta_\psi\\ \widehat{\beta}_1^* - \beta_1^*\\ \widehat{\beta}_2^* - \beta_2^*\end{pmatrix}\xrightarrow{\mathrm{D}}\mathrm{N}\left(0, \mathbb{G}^*\right),$$

i.e. the vector has asymptotically normal distribution with a zero mean vector and a variance - covariance matrix $\mathbb{G}^*$, where $\mathbb{G}^*$ is the inverse matrix of matrix

$$\mathbb{G}^{*-1} = \begin{pmatrix}\beta_1^{*2}\theta_\psi + 2\beta_1^*\beta_2^*\theta_\psi^2 + 4\beta_2^{*2}\theta_\psi^3/3 & \ldots & \ldots\\ \beta_1\theta_\psi^2/2 + 2\beta_2^*\theta_\psi^3/3 & \theta_\psi^3/3 & \ldots\\ \beta_1^*\theta_\psi^3/3 + \beta_2^*\theta_\psi^4/2 & \theta_\psi^4/4 & \theta_\psi^5/5\end{pmatrix}.$$

By inverting the matrix, we calculate

$$\mathbb{G}^* = \begin{pmatrix}\frac{9}{\beta_1^{*2}\theta_\psi} & \ldots & \ldots\\ -\frac{36\beta_1^* - 18\beta_2^*\theta_\psi}{\beta_1^{*2}\theta_\psi^2} & \frac{36\beta_2^{*2}\theta_\psi^2 + 144\beta_1^*\beta_2^*\theta_\psi + 192\beta_1^{*2}}{\beta_1^{*2}\theta_\psi^3} & \ldots\\ \frac{30\beta_1^*}{\beta_1^{*2}\theta_\psi^3} & -\frac{180\beta_1^{*2}\theta_\psi + 60\beta_1^*\beta_2^*\theta_\psi^2}{\beta_1^{*2}\theta_\psi^5} & \frac{180}{\theta_\psi^5}\end{pmatrix}.$$

We need the asymptotic distribution for the estimators $\theta_\psi, \beta_1, \beta_2$ from our original model (2.11). Define a linear transformation

$$g\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\\sigma y\\\sigma z\end{pmatrix}.$$

Since g is continuous, we obtain by using continuous mapping theorem (Theorem 2.3 in Van der Vaart [1998])

$$\sqrt{n}\left(g\begin{pmatrix}\widehat{\theta}_\psi\\\widehat{\beta}_1^*\\\widehat{\beta}_2^*\end{pmatrix}-g\begin{pmatrix}\theta_\psi\\\beta_1^*\\\beta_2^*\end{pmatrix}\right)\xrightarrow{\mathrm{D}}\mathrm{N}\left(0,\mathbb{D}_g\mathbb{G}^*\mathbb{D}_g^\top\right),$$

where $\mathbb{D}_g$ is a the transformation matrix. In our case

$$\mathbb{D}_g=\begin{pmatrix}1&0&0\\0&\sigma&0\\0&0&\sigma\end{pmatrix}.$$

Denote $\mathbb{G}=\mathbb{D}_g^\top\mathbb{G}^*\mathbb{D}_g$. Using $\beta_i^*=\beta_i/\sigma$, the matrix $\mathbb{G}$ equals

$$\mathbb{G}=\begin{pmatrix}\frac{9}{{\beta_1^*}^2\,\theta_\psi}&\cdots&\cdots\\-\frac{\left(36\,\beta_1^*-18\,\beta_2^*\,\theta_\psi\right)\sigma}{{\beta_1^*}^2\,\theta_\psi^2}&\frac{\left(36\,{\beta_2^*}^2\,\theta_\psi^2+144\,\beta_1^*\,\beta_2^*\,\theta_\psi+192\,{\beta_1^*}^2\right)\sigma^2}{{\beta_1^*}^2\,\theta_\psi^3}&\cdots\\\frac{30\,\beta_1^*\,\sigma}{{\beta_1^*}^2\,\theta_\psi^3}&-\frac{\left(180\,{\beta_1^*}^2\,\theta_\psi+60\,\beta_1^*\,\beta_2^*\,\theta_\psi^2\right)\sigma^2}{{\beta_1^*}^2\,\theta_\psi^5}&\frac{180\,\sigma^2}{\theta_\psi^5}\end{pmatrix}$$

$$=\begin{pmatrix}\frac{9\,\sigma^2}{\beta_1^2\,\theta_\psi}&\cdots&\cdots\\-\frac{\left(36\,\beta_1-18\,\beta_2\,\theta_\psi\right)\sigma^2}{\beta_1^2\,\theta_\psi^2}&\frac{\left(36\,\beta_2^2\,\theta_\psi^2+144\,\beta_1\,\beta_2\,\theta_\psi+192\,\beta_1^2\right)\sigma^2}{\beta_1^2\,\theta_\psi^3}&\cdots\\\frac{30\,\beta_1\,\sigma^2}{\beta_1^2\,\theta_\psi^3}&-\frac{\left(180\,\beta_1^2\,\theta_\psi+60\,\beta_1\,\beta_2\,\theta_\psi^2\right)\sigma^2}{\beta_1^2\,\theta_\psi^5}&\frac{180\,\sigma^2}{\theta_\psi^5}\end{pmatrix}.$$

□

Especially we obtain from Theorem 2 the asymptotic marginal distributions

$$\sqrt{n}\left(\widehat{\theta}_\psi-\theta_\psi\right)\sqrt{\frac{\beta_1^2\theta_\psi}{9\sigma^2}}\xrightarrow{\mathrm{D}}\mathrm{N}(0,1)$$

$$\sqrt{n}\,\frac{\widehat{\beta}_1-\beta_1}{\sqrt{v_{\beta_1}}}\xrightarrow{\mathrm{D}}\mathrm{N}(0,1)$$

$$\sqrt{n}\left(\widehat{\beta}_2-\beta_2\right)\sqrt{\frac{\theta_\psi^5}{180\sigma^2}}\xrightarrow{\mathrm{D}}\mathrm{N}(0,1),$$

while denoting $v_{\beta_1} = \frac{\left(36\beta_2^2\theta_\psi^2 + 144\beta_1\beta_2\theta_\psi + 192\beta_1^2\right)\sigma^2}{\beta_1^2\theta_\psi^3}$.

Since σ^2 is unknown, we estimate it by $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n \left(Z_i - \widehat{\beta}_1 x_{i,\widehat{\kappa}} - \widehat{\beta}_2 x_{i,\widehat{\kappa}}^2\right)^2$. Using Cramer-Slutsky theorem (Theorem B.10 in Anděl [2007]) we have

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)\sqrt{\frac{\beta_1^2\theta_\psi}{9\widehat{\sigma}^2}} = \sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)\sqrt{\frac{\beta_1^2\theta_\psi}{9\sigma^2}}\frac{\sigma}{\widehat{\sigma}} \xrightarrow{\mathrm{D}} \mathrm{N}\left(0,1\right),$$

using $\frac{\sigma^2}{\widehat{\sigma}^2} \xrightarrow{\mathrm{P}} 1$ as $n \to \infty$. Using analogous argumentation we obtain

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)\sqrt{\frac{\widehat{\beta}_1^2\widehat{\theta}_\psi}{9\widehat{\sigma}^2}} \xrightarrow{\mathrm{D}} \mathrm{N}\left(0,1\right)$$

$$\sqrt{n}\,\frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\widehat{v}_{\beta_1}}} \xrightarrow{\mathrm{D}} \mathrm{N}\left(0,1\right)$$

$$\sqrt{n}\left(\widehat{\beta}_2 - \beta_2\right)\sqrt{\frac{\widehat{\theta}_\psi^5}{180\widehat{\sigma}^2}} \xrightarrow{\mathrm{D}} \mathrm{N}\left(0,1\right),$$

where $\widehat{v}_{\beta_1} = \frac{\left(36\widehat{\beta}_2^2\widehat{\theta}_\psi^2 + 144\widehat{\beta}_1\widehat{\beta}_2\widehat{\theta}_\psi + 192\widehat{\beta}_1^2\right)\widehat{\sigma}^2}{\widehat{\beta}_1{}^2\widehat{\theta}_\psi^3}$. Using the definition of θ_ψ we have

$$\frac{\widehat{\psi} - \psi}{\sqrt{n}}\sqrt{\frac{\widehat{\beta}_1^2\widehat{\theta}_\psi}{9\widehat{\sigma}^2}} \xrightarrow{\mathrm{D}} \mathrm{N}\left(0,1\right).$$

Next theorem gives the asymptotic distribution when $\beta_1 = 0$.

Theorem 3. *Suppose model* (2.11) *holds,* $\beta_1 = 0$ *and* $\beta_2 \neq 0$. *Let* $(U_1, U_2, U_3)^\top$ *be a normal vector with zero mean and variance-covariance matrix*

$$\begin{pmatrix} \theta_\psi & \theta_\psi^2/2 & \theta_\psi^3/3 \\ \theta_\psi^2/2 & \theta_\psi^3/3 & \theta_\psi^4/4 \\ \theta_\psi^3/3 & \theta_\psi^4/4 & \theta_\psi^5/5 \end{pmatrix}.$$

Let us introduce a random variable $X = -U_1 + \frac{4}{\psi}U_2 - \frac{10}{3\psi^2}U_3$ *having a normal distribution* $\mathrm{N}(0, \theta_\psi/9)$ *and* $U_+ = \max(0, X/\beta_2)$. *Then as* $n \to \infty$

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)^2 \xrightarrow{D} \frac{9}{\theta_\psi}\,U_+ \equiv \max\left(0, \mathrm{N}\left(0, \frac{9\sigma^2}{\beta_2{}^2\theta_\psi}\right)\right),$$

$$\sqrt{n}\,\widehat{\beta}_1^2 \xrightarrow{D} \frac{36\sigma}{\theta_\psi}\beta_2{}^2\,U_+ \equiv \max\left(0, \mathrm{N}\left(0, \frac{144\beta_2^2\sigma^2}{\theta_\psi}\right)\right),$$

$$\sqrt{n}\,\left(\widehat{\beta}_2 - \beta_2\right) \xrightarrow{D} -\frac{30\sigma}{\theta_\psi^3}\beta_2 U_+ - \frac{60\sigma}{\theta_\psi^4}U_2 + \frac{80\sigma}{\theta_\psi^5}U_3.$$

Proof. We will proceed similarly as in Theorem 2 and we will use Theorem B from Jarušková [2001]. Define $Z_i^* = \frac{Z_i}{\sigma}$, it holds

$$Z_i^* = \frac{\beta_2}{\sigma}\left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + \frac{e_i}{\sigma} = \beta_2^*\left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + e_i^*$$

denoting $\beta_i^* = \beta_i/\sigma$ and $e_i^* = e_i/\sigma$. From the data $Z_1^*, \ldots, Z_n^*$ we obtain estimators $\widehat{\theta}_\psi$, $\widehat{\beta}_1^*$, $\widehat{\beta}_2^*$. The data $Z_1^*, \ldots, Z_n^*$ satisfy the model in Jarušková [2001]. Define $U_+^* = \max(0, X/\beta_2^*) = \sigma U_+$ by using Theorem B in Jarušková [2001] we obtain

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)^2 \xrightarrow{\text{D}} \frac{9}{\theta_\psi} U_+^* \equiv \max\left(0, \text{N}\left(\frac{9}{\beta_2^{*2}\theta_\psi}\right)\right),$$

$$\sqrt{n}\,\widehat{\beta}_1^{*2} \xrightarrow{\text{D}} \frac{36}{\theta_\psi}\beta_2^{*2}\, U_+^* \equiv \max\left(0, \text{N}\left(\frac{144\beta_2^{*2}}{\theta_\psi}\right)\right),$$

$$\sqrt{n}\left(\widehat{\beta}_2^* - \beta_2^*\right) \xrightarrow{\text{D}} -\frac{30}{\theta_\psi^3}\beta_2^* U_+^* - \frac{60}{\theta_\psi^4}U_2 + \frac{80}{\theta_\psi^5}U_3\,.$$

Since the change-point ψ is the same in data $Z_1, \ldots, Z_n$ and $Z_1^*, \ldots, Z_n^*$, by using $\beta_i^* = \beta_i/\sigma$ we obtain

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)^2 \xrightarrow{\text{D}} \frac{9}{\theta_\psi}\sigma\, U_+ \equiv \max\left(0, \text{N}\left(\frac{9\sigma^2}{{\beta_2}^2\theta_\psi}\right)\right).$$

Using a linear transformation $g(x) = \sigma^2 x$ and continuous mapping theorem (Theorem 2.3 in Van der Vaart [1998]) we obtain, as $n \to \infty$,

$$\sqrt{n}\,\widehat{\beta}_1^{*2}\sigma^2 = \sqrt{n}\,\widehat{\beta}_1^2 \xrightarrow{\text{D}} \frac{36\sigma^2}{\theta_\psi}\beta_2^{*2}\, U_+^* = \frac{36}{\theta_\psi}{\beta_2}^2\sigma\, U_+ \equiv \max\left(0, \text{N}\left(\frac{144\beta_2^2\sigma^2}{\theta_\psi}\right)\right).$$

By using a linear transformation $g(x) = \sigma x$ and continuous mapping theorem, we obtain

$$\sqrt{n}\,\sigma\left(\widehat{\beta}_2^* - \beta_2^*\right) = \sqrt{n}\left(\widehat{\beta}_2 - \beta_2\right) \xrightarrow{\text{D}} -\frac{30}{\theta_\psi^3}\beta_2\sigma U_+ - \frac{60\sigma}{\theta_\psi^4}U_2 + \frac{80\sigma}{\theta_\psi^5}U_3.$$

□

The asymptotic results can be used to construct the confidence intervals for unknown parameters. The two-sided (5.1) and right-sided (5.2) confidence intervals for the change-point ψ based on Theorem 2 are stated in Chapter 5 and their coverage is examined in simulations.

2.2.2 Unknown β_0

With unknown β_0, we have model (2.10). The estimation of unknown parameters can be done similarly as in Chapter 2 with $d = 2$.

Keeping the notation $x_{p,i}^s = \left(\left(\frac{\psi - i}{n}\right)^+\right)^s$ for $s = 1, 2$, we have

$$\mathbb{X}_{p\cdot} = \begin{pmatrix} 1 & x_{p,1} & x_{p,1}^2 \\ 1 & x_{p,2} & x_{p,2}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{p,n} & x_{p,n}^2 \end{pmatrix}.$$

The change-point ψ can be estimated similarly as in (2.2) by

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max}\ \boldsymbol{Z}^\top \mathbb{X}_{p\cdot} \left(\mathbb{X}_{p\cdot}^\top \mathbb{X}_{p\cdot}\right)^{-1} \mathbb{X}_{p\cdot}^\top \boldsymbol{Z}.$$

According to remark from the first Section, we can equivalently estimate ψ by

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max}\ R_p^2, \tag{2.13}$$

where R_p^2 denotes the coefficient of determination of the linear regression model with response $\boldsymbol{Z}$ and model matrix $\mathbb{X}_{p\cdot}$.

Parameters $\boldsymbol{\beta} = (\beta_0,\ \beta_1,\ \beta_2)^\top$ can then be estimated as in (2.4), i.e.

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\widehat{\psi}\cdot}^\top\, \mathbb{X}_{\widehat{\psi}\cdot}\right)^{-1} \mathbb{X}_{\widehat{\psi}\cdot}^\top\, \boldsymbol{Z}\,. \tag{2.14}$$

With model (2.10) we do not have any other simpler formula of $\widehat{\boldsymbol{\beta}}, \widehat{\psi}$ as in (2.8) since the matrix $\mathbb{X}_{p\cdot}$ is more complicated in this case.

Denote $\theta_\psi = \psi/n$. Assuming $\operatorname{var} e_i = 1$, it can be proved, that as $n \to \infty$ and $\beta_1 \neq 0$

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right) \xrightarrow{\ \mathrm{D}\ } \mathrm{N}\left(0, \frac{9 - 5\theta_\psi}{\beta_1^2 \theta_\psi \left(1 - \theta_\psi\right)}\right)$$

and if $\beta_1 = 0$, $\beta_2 \neq 0$ and $n \to \infty$

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)^2 \xrightarrow{\ \mathrm{D}\ } \max\left(0, \mathrm{N}\left(0, \frac{9 - 5\theta_\psi}{\beta_2^2 \theta_\psi \left(1 - \theta_\psi\right)}\right)\right),$$

using Remark 3 in Jarušková [2001] and arguments as in Theorem 4 with „reversed“ time ordering.

Similarly as in Theorem 2 we can modify the results for the case with $\sigma^2 \neq 1$. Suppose $\beta_1 \neq 0$ and define $Z_i^* = Z_i/\sigma$. By using Remark 3 from Jarušková [2001] for data $Z_1^*, \ldots, Z_n^*$ we obtain

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right) \xrightarrow{\ \mathrm{D}\ } \mathrm{N}\left(0, \frac{\left(9 - 5\theta_\psi\right)\sigma^2}{\beta_1^2 \theta_\psi \left(1 - \theta_\psi\right)}\right). \tag{2.15}$$

If $\beta_1 = 0$ and $\beta_2 \neq 0$ we have

$$\sqrt{n}\left(\widehat{\theta}_\psi - \theta_\psi\right)^2 \xrightarrow{\ \mathrm{D}\ } \max\left(0, \mathrm{N}\left(0, \frac{\left(9 - 5\theta_\psi\right)\sigma^2}{\beta_2^2 \theta_\psi \left(1 - \theta_\psi\right)}\right)\right),$$

2.3 E_{max} model

We now focus on E_{max} model, which is often used in dose-response studies where it models the output with respect to patients with different doses, see e.g. MacDougall [2006], but it can be used also in other contexts.

Let us have data $Z_1, \ldots, Z_n$ measured at times $i = 1, \ldots, n$ from the E_{max} model

$$
\begin{aligned}
Z_i &= \beta_{lowAs} + \beta_{incr} \frac{i}{i+\gamma} + e_i \\
&= \beta_{lowAs} + \beta_{incr} \frac{1}{1+\frac{\gamma}{i}} + e_i
\end{aligned}
$$

where β_{lowAs} is the baseline (the lower asymptote), β_{incr} represents the effect and γ denotes the time when the effect is one half. Random errors are assumed to have zero expectation and variance σ^2. With $\beta_{incr} > 0$ the expected value of Y_i increases from β_{lowAs} towards $\beta_{lowAs} + \beta_{incr}$.

Now, we include a change-point ψ to the E_{max} model. For time i before change-point ψ we have

$$
Z_i = \beta_{lowAs} + \beta_{incr} \frac{1}{1+\frac{\gamma}{i}} + e_i
$$

and for $i \geq \psi$, we have

$$
Z_i = \beta_{lowAs} + \beta_{incr} \frac{1}{1+\frac{\gamma}{\psi}} + e_i.
$$

We can rewrite these formulae together as

$$
Z_i = \beta_{lowAs} + \beta_{incr} \frac{1}{1+\frac{\gamma}{\min(i,\psi)}} + e_i. \tag{2.16}
$$

Before change-point, the data follow the E_{max} model and after ψ the expected value of Z_i remains the same as for Z_ψ, see Figure 2.1.

We focus on the E_{max} model with a change-point. The parameters of the model are ψ, γ, σ^2 and $\boldsymbol{\beta} = (\beta_{lowAs}, \beta_{incr})^\top$. The estimation proceeds similarly as in polynomial gradual change model, we only have more general $x_{p,i}$. Denote

$$
x_{p,i;\,c} = \frac{1}{1+\frac{c}{\min(p,i)}}
$$

and

$$
\mathbb{X}_{p,\cdot;\,c} = \begin{pmatrix} 1 & x_{p,1;\,c} \\ \vdots & \vdots \\ 1 & x_{p,n;\,c} \end{pmatrix}.
$$

The model (2.16) can be written as $\boldsymbol{Y} = \mathbb{X}_{\psi,\cdot;\,\gamma}\, \boldsymbol{\beta} + \boldsymbol{e}$ where $\boldsymbol{e} = (e_1, \ldots, e_n)^\top$, therefore we can use formulae from Section 2 to estimate the parameters, we only have more general model matrix $\mathbb{X}$.

During the estimation, we solve

$$\min_{\substack{p,\,c\,\in\,(1,n)\\ \boldsymbol{\beta}\,\in\,\mathbb{R}}} \|\boldsymbol{Y} - \mathbb{X}_{p,\cdot;\,c}\boldsymbol{\beta}\| = \min_{\substack{p,\,c\,\in\,(1,n)\\ \boldsymbol{\beta}\,\in\,\mathbb{R}}} \left(\boldsymbol{Y} - \mathbb{X}_{p,\cdot;\,c}\boldsymbol{\beta}\right)^\top \left(\boldsymbol{Y} - \mathbb{X}_{p,\cdot;\,c}\boldsymbol{\beta}\right).$$

Similarly as in (2.2) we obtain

$$\begin{aligned}\widehat{\psi},\ \widehat{\gamma} &= \underset{p,\,c\,\in\,(1,n)}{\arg\min}\ \boldsymbol{Y}^\top \left(\mathbb{I} - \mathbb{X}_{p,\cdot;\,c}\left(\mathbb{X}_{p,\cdot;\,c}^\top \mathbb{X}_{p,\cdot;\,c}\right)^{-1} \mathbb{X}_{p,\cdot;\,c}^\top\right) \boldsymbol{Y} \\ &= \underset{p,\,c\,\in\,(1,n)}{\arg\max}\ \boldsymbol{Y}^\top \mathbb{X}_{p,\cdot;\,c}\left(\mathbb{X}_{p,\cdot;\,c}^\top \mathbb{X}_{p,\cdot;\,c}\right)^{-1} \mathbb{X}_{p,\cdot;\,c}^\top \boldsymbol{Y},\end{aligned} \tag{2.17}$$

where the minimisation is done over two parameters, p and c. With estimated change-point ψ and the time of the half effect γ, we can estimate parameters $\boldsymbol{\beta} = (\beta_{lowAs},\ \beta_{incr})^\top$ by

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\widehat{\psi},\cdot;\,\widehat{\gamma}}^\top \mathbb{X}_{\widehat{\psi},\cdot;\,\widehat{\gamma}}\right)^{-1} \mathbb{X}_{\widehat{\psi},\cdot;\,\widehat{\gamma}}^\top \boldsymbol{Y} \tag{2.18}$$

and σ^2 by

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{\beta}_{lowAs} - \widehat{\beta}_{incr} x_{i;\widehat{\psi},\widehat{\gamma}}\right)^2. \tag{2.19}$$

The distribution of derived estimators can be approximated using bootstrap, the procedure will be described in more detail in Chapter 5.

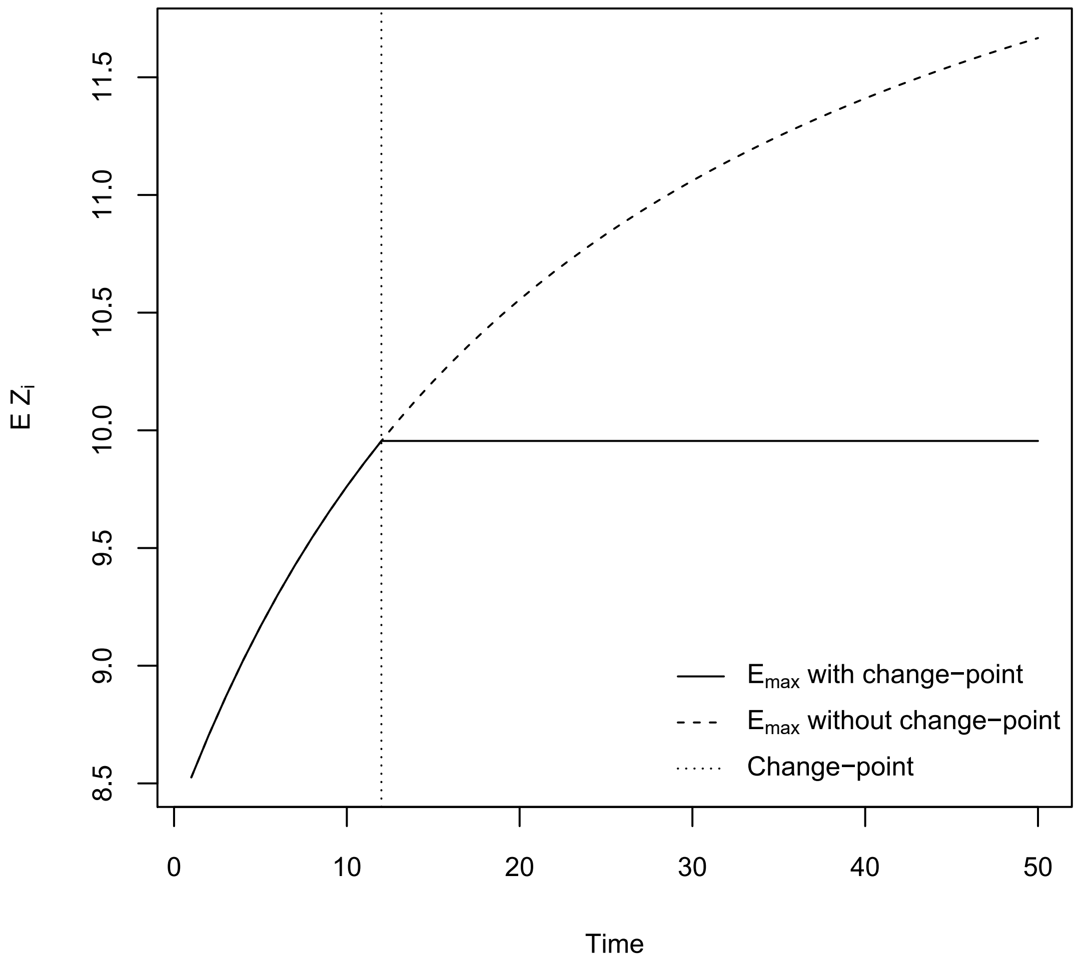

Figure 2.1: Expected value of E_{max} model with and without change-point

3. Linear point-of-stabilisation model

In this chapter, we work with the gradual change model with linear trend. Moreover, we modify the model to fit into context of so-called *point-of-stabilisation* (*PoSt*), which was mentioned in Section 2.2.

In drug manufacturing process, the continuous manufacturing started replacing batch process because of its better properties, described e.g. in Lee et al. [2015] and in Schaber et al. [2011]. In continuous manufacturing, the product is moving through the production line and mixing as applied continuously instead of applying it to whole batch in batch process. Continuous manufacturing is more agile and robust, since often in classical batch process manufacturing, the outputs from one step of production are tested, stored and transported to the next step, which takes some time and increases costs. On the other hand, continuous manufacturing eliminates these factors as the outputs of one step are tested and immediately transported to the next step. It also shortens the supply chain since in batch process manufacturing, the materials can be shipped across several countries before they are applied in next step of the process, see e.g. Lee et al. [2015].

Continuous manufacturing can be scaled-up in a easier way than batch manufacturing either by increasing flow rate of the line or by creating a new one, and can therefore prevent drug shortage in case of unforeseen situations. The cost effectiveness was studied e.g. in Schaber et al. [2011] and it turned out that continuous manufacturing brings savings compared to the batch process, especially when more sophisticated processes of continuous manufacturing are used.

During the start-up period of the production line, the final products do not have the same quality. This brings material waste during this starting phase of manufacturing.

The goal is to find the so-called *point-of-stabilisation* (*PoSt*), when the process becomes stable and it stops showing the trend observed in early phase. This problem can be viewed as a simple modification of the change-point estimation in gradually changing sequence with monotonic trend. Point estimate as well as confidence interval (usually right-sided) for change-point is desired.

Here, we suppose the linear trend to be present up to an unknown change-point ψ (the *PoSt*) and diminish after ψ, similarly as in Section 2.2. In continuous drug manufacturing, the part before ψ represents the start-up period with unstable output quality. Assume we have data $Z_1, \ldots, Z_n$ satisfying

$$Z_i = \beta_0 + \beta_1 \left(\frac{\psi - i}{n} \right)^+ + e_i. \tag{3.1}$$

Random errors $e_1, \ldots, e_n$ are supposed to be as in model (1.1). Model (3.1) is similar to model (1.1) for $d = 1$ but the time is in reversed order, see Figures 3.3 and 1.1. Also, the expected value the output variable Z is often increasing before the change-point by having $\beta_1 < 0$. After the process becomes stable, i.e. after the unknown time ψ, the output quality of the drug product (e.g. a tablet) does not show any trend. Therefore, in the *point-of-stabilisation* context, we use notation as in Section 2.2 and 2.3 where the trend was also assumed to be present up to the change-point ψ.

3.1 Testing in PoSt model

The main focus is on the estimation, therefore we just briefly discuss testing in PoSt model (3.1) and differences compared to models from previous chapters. Similarly as in Chapter 1, tests about the change-point (here denoted ψ) can be developed. But in model (3.1), the situation is slightly different because the trend appears at the beginning and diminishes afterwards.

Under the null hypothesis $H_0 : \psi = n$, the trend is present in all the data and there is no change-point while at Chapter 1, $H_0 : \kappa = n$ corresponded to a situation without any trend and constant model. Under the alternative $H_1 : \psi < n$, the trend diminishes after ψ and the change is present in our data. Again, in Chapter 1, H_1 corresponded to situation when the trend appeared after the change-point.

3.2 Estimation in PoSt model

We will derive the estimates of unknown parameters in model (3.1) and their asymptotic distribution as in Hušková [1998] while taking into account the reversed time ordering.

As in Section 2.2, denote $x_{p,i} = ((p-i)/n)^+$ and $\overline{x}_{p\cdot} = \left(\sum_{i=1}^n x_{p,i}\right)/n$. The matrix $\mathbb{X}_{p\cdot}$ has form

$$\mathbb{X}_{p\cdot} = \begin{pmatrix} 1 & x_{p,1} \\ 1 & x_{p,2} \\ \vdots & \vdots \\ 1 & x_{p,n} \end{pmatrix}.$$

In this situation, the unknown change-point ψ can be estimated similarly as in (2.2). Denoting $\boldsymbol{Z} = (Z_1, \ldots, Z_n)^\top$ and assuming model (3.1), we have for each $p \in (1, n)$

$$\boldsymbol{Z}^\top \boldsymbol{Z} - \boldsymbol{Z}^\top \mathbb{X}_{p\cdot} \left(\mathbb{X}_{p\cdot}^\top \mathbb{X}_{p\cdot}\right)^{-1} \mathbb{X}_{p\cdot}^\top \boldsymbol{Z} = \sum_{i=1}^n \left(Z_i - \overline{Z}\right)^2 - \frac{\sum_{i=1}^n Z_i \left(x_{p,i} - \overline{x}_{p\cdot}\right)^2}{\sum_{i=1}^n \left(x_{p,i} - \overline{x}_{p\cdot}\right)^2}.$$

Therefore the estimator $\widehat{\psi}$ of the change-point ψ can be calculated as in (2.2) or equivalently as

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max} \frac{\left(\sum_{i=1}^n Z_i \left(x_{p,i} - \overline{x}_{p\cdot}\right)\right)^2}{\sum_{i=1}^n \left(x_{p,i} - \overline{x}_{p\cdot}\right)^2}. \tag{3.2}$$

The shape of minimised function depends on the location of ψ, see Figure 3.1 created with $n = 50$, $\beta_0 = \beta_1 = 2$ and $\psi = [n/4]$, $[n/2]$, $[3n/4]$.

Similarly as in (2.4), the estimator $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ has form

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\widehat{\psi}\cdot}^\top \mathbb{X}_{\widehat{\psi}\cdot}\right)^{-1} \mathbb{X}_{\widehat{\psi}\cdot}^\top \boldsymbol{Z},$$

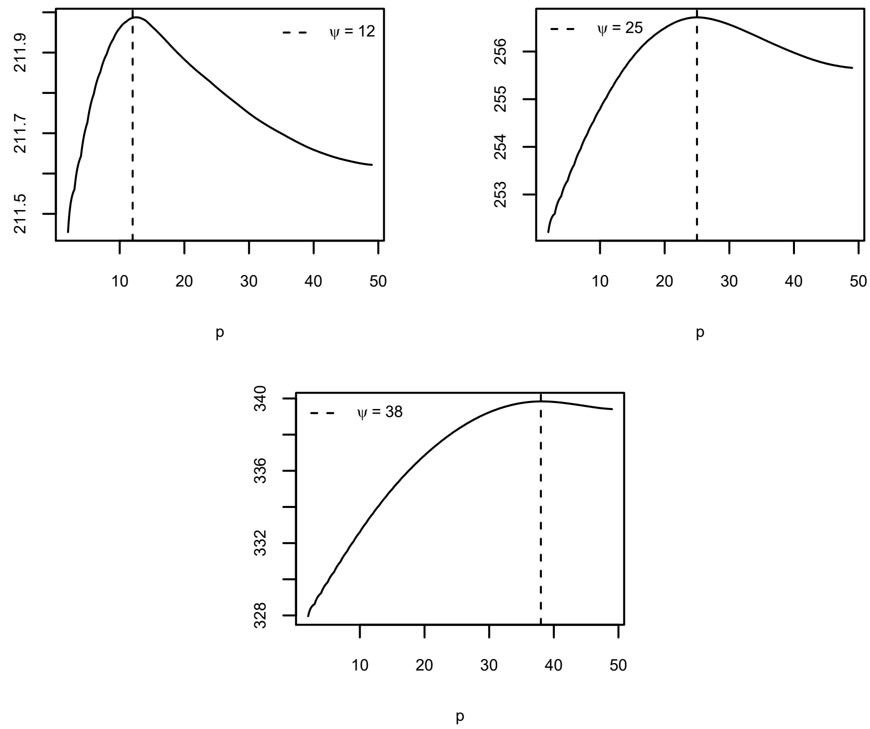

Figure 3.1: Function to be maximised in linear *PoSt* model for $n = 50$, $\beta_0 = \beta_1 = 2$ and $\psi = [n/4]$, $[n/2]$, $[3n/4]$.

specifically in our case

$$\widehat{\beta}_0 = \overline{Z}_n - \widehat{\beta}_1 \, \overline{x}_{\widehat{\psi}\cdot}.$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n Z_i \left(x_{\widehat{\psi},i} - \overline{x}_{\widehat{\psi}\cdot} \right)}{\sum_{i=1}^n \left(x_{\widehat{\psi},i} - \overline{x}_{\widehat{\psi}\cdot} \right)^2}.$$

The variance σ^2 can be estimated similarly as in (2.5) by

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{\widehat{\psi},i} \right)^2 .$$

The asymptotic distribution of the estimators can be derived similarly as in Hušková [1998]. We only have reverse time ordering and therefore slightly different model in our case. We have to modify our data a little bit and then use the theorems from Hušková [1998].

Theorem 4. *Let the variables* $Z_1, \ldots, Z_n$ *satisfy* (3.1) *and be independent. Let, as* $n \to \infty$,

$$\psi = [n\theta_\psi]$$

for some $\theta_\psi \in (0, 1)$.

Then, as $n \to \infty$,

$$\frac{\beta_1}{\sigma} \frac{\widehat{\psi} - \psi}{\sqrt{n}} \sqrt{\frac{\left(1 - \theta_\psi\right) \theta_\psi}{4 - 3\theta_\psi}} \xrightarrow{D} \mathrm{N}(0, 1). \tag{3.3}$$

Proof. We would like to use Theorem 1 from Hušková [1998] but we have to modify our data to fit model (1.1).

Define $Y_i = Z_{n-i+1}$ for $i = 1, \dots, n$. Then

$$Z_{n-i+1} \;=\; \beta_0 + \beta_1 \left(\frac{\psi - (n-i+1)}{n} \right)^+ + e_{n-i+1} \;=\; \beta_0 + \beta_1 \left(\frac{i-\kappa}{n} \right)^+ + \widetilde{e}_i,$$

where $\kappa = n - \psi + 1$ and $\widetilde{e}_i = e_{n-i+1}$. The data $Y_i, \dots, Y_n$ satisfy a model

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 \left(\frac{i-\kappa}{n} \right)^+ + \widetilde{e}_i \\ &= \beta_0 + \beta_1 x_{i,\kappa} + \widetilde{e}_i, \end{aligned}$$

where β_0, β_1 are as in (3.1). This is like the original model (3.1) with „reversed“ time and new change-point $\kappa = n - \psi + 1$. We are going to apply the theorem to data $Y_1, \dots, Y_n$ to derive the asymptotic distribution of the estimator of κ. Then we show it can be used also for the estimator of ψ.

The estimators of β_0, β_1, σ^2 given by (2.8) and (2.2) calculated from $Y_1, \dots, Y_n$. are the same as if calculated using $Z_1, \dots, Z_n$. This can be seen from formulae (2.8), (2.2) and the fact that $x_{i,k} = x_{p,n-i+1}$ for each i, k and $p = n - k + 1$.

The change-point κ can be estimated as

$$\widehat{\kappa} = \operatorname*{arg\,max}_{k \in (1,n)} \frac{\left(\sum_{i=1}^n Y_i \left(x_{i,k} - \overline{x}_{\cdot k} \right) \right)^2}{\sum_{i=1}^n \left(x_{i,k} - \overline{x}_{\cdot k} \right)^2} \tag{3.4}$$

and it holds $\widehat{\kappa} = n - \widehat{\psi} + 1$.

Further, $\theta_\psi = \frac{\psi}{n} = \frac{n-\kappa+1}{n} = \frac{n+1}{n} - \theta_\kappa$ while denoting $\theta_\kappa = \frac{k}{n}$. The parameter β_1 is assumed not to depend on n, therefore as $n \to \infty$

$$\beta_1 = O(1), \qquad \frac{\beta_1^2 n}{(\log \log n)} \to \infty\,.$$

From Theorem 1 we have, as $n \to \infty$,

$$\frac{\beta_1}{\sigma} \frac{\widehat{\kappa} - \kappa}{\sqrt{n}} \sqrt{\frac{\theta_\kappa (1-\theta_\kappa)}{1+3\theta_\kappa}} \xrightarrow{\text{D}} \mathrm{N}(0,1).$$

Since $\widehat{\kappa} - \kappa = \psi - \widehat{\psi}$ and $\theta_\kappa = \frac{n+1}{n} - \theta_\psi \to 1 - \theta_\psi$ as $n \to \infty$, we have as $n \longrightarrow \infty$,

$$\frac{\beta_1}{\sigma} \frac{\widehat{\psi} - \psi}{\sqrt{n}} \sqrt{\frac{\left(1-\theta_\psi\right)\theta_\psi}{4-3\theta_\psi}} \xrightarrow{\text{D}} \mathrm{N}(0,1).$$

□

Similarly, we obtain the asymptotic distribution of the other estimators using results from Hušková [1998].

Theorem 5. *Let the assumptions of Theorem 4 be satisfied. Then, as $n \to \infty$,*

$$\sqrt{n}\left(\widehat{\beta}_1 - \beta_1\right) \xrightarrow{D} \mathrm{N}\left(0, \frac{12\sigma^2}{\theta_\psi^3\left(4 - 3\theta_\psi\right)}\right),$$

$$\sqrt{n}\left(\widehat{\beta}_0 - \beta_0\right) \xrightarrow{D} \mathrm{N}\left(0, \frac{4\sigma^2}{4 - 3\theta_\psi}\right),$$

$$\widehat{\sigma}^2 - \sigma^2 = o_P\left((\log\log n)^{-1}\right),$$

Proof. Hušková [1998, Theorem B].

□

We will now examine the results of Theorem 4 in more detail. Denote

$$V(\theta) = \sqrt{\frac{1+3\theta}{\theta\,(1-\theta)}}.$$

We can rewrite the results of Theorem 1 (model (1.1) with $d = 1$) as

$$\frac{\beta_1}{\sigma}\frac{\widehat{\kappa} - \kappa}{\sqrt{n}\,V(\theta)} \xrightarrow{\mathrm{D}} \mathrm{N}(0, 1),$$

where $\theta = \kappa/n$. The term $V(\theta)$ can be consistently estimated by $V(\widehat{\theta})$. The expression $V(\theta)$ is part of the variance of $\widehat{\theta} - \theta$ and changes with changing θ.

In model (3.1) with change-point ψ (representing PoSt) we work with „reverse" time ordering. To be able to transform the data $Z_1, \ldots, Z_n$ to data satisfying Theorem 1 we introduced „reversed" change-point $\kappa = n - \psi + 1$. We can write $\theta_\psi = \psi/n = (n+1)/n - \theta_\kappa$, where $\theta_\kappa = \kappa/n$ and $\theta_\psi \to 1 - \theta_\kappa$ for $n \to \infty$. The asymptotic distribution of $\widehat{\psi}$ in Theorem 4 can be written as

$$\frac{\beta_1}{\sigma}\frac{\widehat{\psi} - \psi}{\sqrt{n}V(1-\theta_\psi)} \xrightarrow{\mathrm{D}} \mathrm{N}(0, 1)$$

where $V(1-\theta_\psi) = \sqrt{\frac{4-3\theta_\psi}{\left(1-\theta_\psi\right)\theta_\psi}}$. We see values of $V(1-\theta_\psi)$ as a function of θ_ψ in Figure 3.2. The minimum lies at $\theta = 2/3$ and values increases for smaller or bigger θ_ψ. The highest values of $V(1-\theta_\psi)$ are for small θ_ψ. One could expect such behaviour as this represents the situation when we do not have enough data before change.

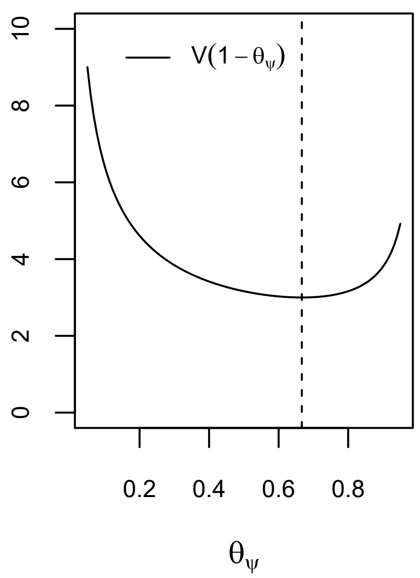

Figure 3.2: Values of $V(1-\theta_\psi) = \sqrt{\frac{4-3\theta_\psi}{(1-\theta_\psi)\theta_\psi}}$ as a function of θ_ψ in linear *PoSt* model (3.1).

3.3 Confidence intervals

The estimator $\widehat{\psi}$ of the *point-of-stabilisation* ψ is given by (3.2). Theorem 4 can then be used to construct confidence intervals for the change-point ψ. Often, one-sided confidence interval for ψ is desired because of its equivalence to hypothesis testing. In practice, we are often facing a one-sided problem because we want to be sure the stabilisation was achieved and the quality of the product is guaranteed. In that case, we want to verify the stability of the process at given time ψ_0, i.e. that no trend is present at ψ_0. The trend is assumed to be caused by the start-up period of the production line, which stabilises at time ψ. Therefore we want to test, whether at given time ψ_0, the output quality of products is without any trend or equivalently whether $\psi < \psi_0$, because the trend is present up to ψ, see Figure 3.3. We put the desired case as the alternative and we obtain

$$H_0 : \psi \geq \psi_0$$
$$H_1 : \psi < \psi_0\,.$$

Using the results of Theorem 4 we reject H_0 if and only if

$$\frac{\beta_1}{\sigma}\frac{\widehat{\psi}-\psi_0}{\sqrt{n}}\sqrt{\frac{\left(1-\widehat{\theta}_\psi\right)\widehat{\theta}_\psi}{4-3\widehat{\theta}_\psi}} \leq u_\alpha$$

where u_α denotes the α-quantile of standard normal distribution. From duality of testing and confidence intervals, the set of values θ_0 for which we do not reject

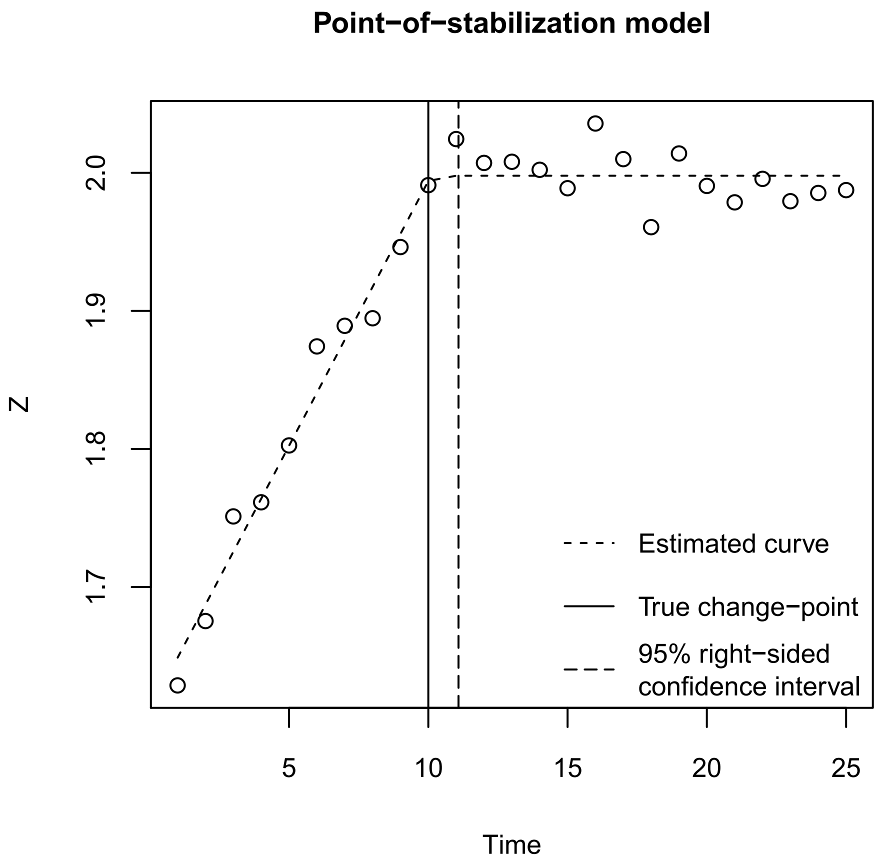

Figure 3.3: Linear *PoSt* model with the right-sided 95% confidence interval for *point-of-stabilisation* ψ given by (3.5).

H_0 with given data represents (if they form an interval) the confidence interval for θ with an asymptotic coverage $1-\alpha$. Since H_0 is not rejected when the test statistic is greater than $u_\alpha = -u_{1-\alpha}$ we obtain an interval

$$(-\infty, c_U) = \left(-\infty,\ \widehat{\psi} + u_{1-\alpha}\frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta}_1}\sqrt{\frac{4-3\widehat{\theta}_\psi}{\widehat{\theta}_\psi\left(1-\widehat{\theta}_\psi\right)}}\right), \tag{3.5}$$

Therefore when we reject H_0 if ψ_0 does not lie in this interval, such test will have an asymptotic significance level α.

On the other hand, we can also test

$$H_0 : \mathsf{E}\ Z_{\psi_0} = \beta_0$$
$$H_1 : \mathsf{E}\ Z_{\psi_0} \neq \beta_0 \,.$$

Under H_0, no trend is present at ψ_0 because the change-point ψ occurred before ψ_0, therefore it is equivalent to testing

$$H_0 : \psi \leq \psi_0$$
$$H_1 : \psi > \psi_0 \,.$$

Under the alternative, the true change-point is located after ψ_0 and therefore the trend is still present at ψ_0. We test these hypothesis when we want to verify the presence of the trend at ψ_0 (the alternative) or equivalently, when we want to verify that the process is unstable. Using the results of Theorem 4 we reject H_0 if and only if

$$\frac{\beta_1}{\sigma}\frac{\widehat{\psi}-\psi_0}{\sqrt{n}}\sqrt{\frac{\left(1-\widehat{\theta}_\psi\right)\widehat{\theta}_\psi}{4-3\widehat{\theta}_\psi}} \geq u_{1-\alpha}$$

where u_α denotes the α-quantile of standard normal distribution. From duality of testing and confidence intervals, the set of values θ_0 for which we do not reject H_0 with given data represents (if they form an interval) the confidence interval for θ with an asymptotic coverage $1-\alpha$. Since H_0 is not rejected when the test statistic is lower than $u_{1-\alpha}$ we obtain an interval

$$(c_L,\infty) = \left(\widehat{\psi} - u_{1-\alpha}\frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta}_1}\sqrt{\frac{4-3\widehat{\theta}_\psi}{\widehat{\theta}_\psi\left(1-\widehat{\theta}_\psi\right)}},\ \infty\right). \tag{3.6}$$

Hence if we construct this interval and we reject the hypothesis of stable process at ψ_0 (no trend present) in favour of present trend (H_1) at time ψ_0 if ψ_0 does not lie in the confidence interval, such test will have an asymptotic significance level α.

Using again the results of Theorem 4, the two-sided confidence interval for ψ has form

$$\left(\widehat{\psi} - u_{1-\frac{\alpha}{2}}\frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta}_1}\sqrt{\frac{4-3\widehat{\theta}_\psi}{\widehat{\theta}_\psi\left(1-\widehat{\theta}_\psi\right)}},\ \widehat{\psi} + u_{1-\frac{\alpha}{2}}\frac{\widehat{\sigma}\sqrt{n}}{\widehat{\beta}_1}\sqrt{\frac{4-3\widehat{\theta}_\psi}{\widehat{\theta}_\psi\left(1-\widehat{\theta}_\psi\right)}}\right). \tag{3.7}$$

3.4 Known β_0

So far, we assumed β_0 to be unknown. There are cases when we can assume β_0 to be known. For example, in *PoSt* model (3.1), β_0 represents the mean value of products output quality after the start-up period and it can assumed to be 1, since the quality is often represented by percentage of given target, e.g. percentage of targeted drug load in one tablet.

Assume β_0 to be known and without loss of generality it can be set to zero. In PoSt model, the β_0 is usually 1 but we can define $\widetilde{Z_i} = Z_i - 1$ and work with data $\widetilde{Z_1},\ldots,\widetilde{Z_n}$. With $\beta_0 = 0$ we get for $p \in (1,n)$

$$\mathbb{X}_{p\cdot} = \begin{pmatrix} x_{p,1} \\ x_{p,2} \\ \vdots \\ x_{p,n} \end{pmatrix}.$$

The estimates of ψ and β_1 have form

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max} \frac{\left(\sum_{i=1}^n Z_i x_{p,i}\right)^2}{\sum_{i=1}^n x_{p,i}^2},$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n Z_i x_{\widehat{\psi},i}}{\sum_{i=1}^n x_{\widehat{\psi},i}^2}.$$

The asymptotic distribution of $\widehat{\psi}$ also changes. Using results from Hlávka and Hušková [2017] and similar arguments as in Theorem 4, it holds, as $n \to \infty$,

$$\frac{\beta_1}{\sigma} \frac{\widehat{\psi} - \psi}{\sqrt{n}} \sqrt{\frac{\theta_\psi}{4}} \xrightarrow{\text{D}} \text{N}(0,1),$$

$$\sqrt{n}\left(\widehat{\beta_1} - \beta_1\right) \sqrt{\frac{\theta_\psi^3}{3\sigma^2}} \xrightarrow{\text{D}} \text{N}(0,1).$$

3.5 Simulations

In this section, we run simulations to validate the asymptotic distribution, we construct confidence intervals based on the asymptotic distribution and bootstrap and we simulate their coverage for different parameter choices.

First, we examine results of Theorem 4 and Theorem 5. To visualise the convergence to the asymptotic distribution with finite sample sizes, we simulate 1000 samples and draw their histogram together with a curve representing the asymptotic distribution.

Properties of estimator $\widehat{\psi}$ depend on σ, location of the true ψ and sample size n. We start with investigating σ, see Figure 3.4. For small σ, the variance of the asymptotic distribution is small and the histogram nicely approximates the density of the asymptotic distribution. For $\sigma = 0.05$, simulated values of $\widehat{\psi}$ are very near the true value $\psi = 25$. With increasing σ, the variance of the asymptotic distribution also increases and the estimators are more dispersed, see Figure 3.4.

We also obtain different results for different locations of the change-points, see Figure 3.5. With change-point located at small time i, estimators are more dispersed than for larger times i. Similar behaviour caused by the shape of the variance as a function of θ_ψ was discussed after Theorem 5, see Figure 3.2. The results are similar for β_0 unknown and known.

For estimators $\widehat{\boldsymbol{\beta}}$, we simulate 1000 samples of $\widehat{\beta}_0$, $\widehat{\beta}_1$ when β_0 is unknown and 1000 samples of $\widehat{\beta}_1$ when β_0 is known and assumed to be zero and we draw their histograms together with the asymptotic distribution for different values of σ, see Figure 3.6 and 3.7. The histograms approximate the asymptotic distribution of β_0 nicely for all selected σ but the situation is worse for β_1, see Figure 3.6. Simulated values have higher variance than the variance of the asymptotic distribution. For $\sigma = 0.3$, some of the simulated $\widehat{\beta_1}$ differ from true β_1 by more than one. With known β_0, the situation is similar, see Figure 3.7. The simulated values of $\widehat{\beta}_1$ are again more dispersed than they should be according to the asymptotic distribution, especially for larger σ.

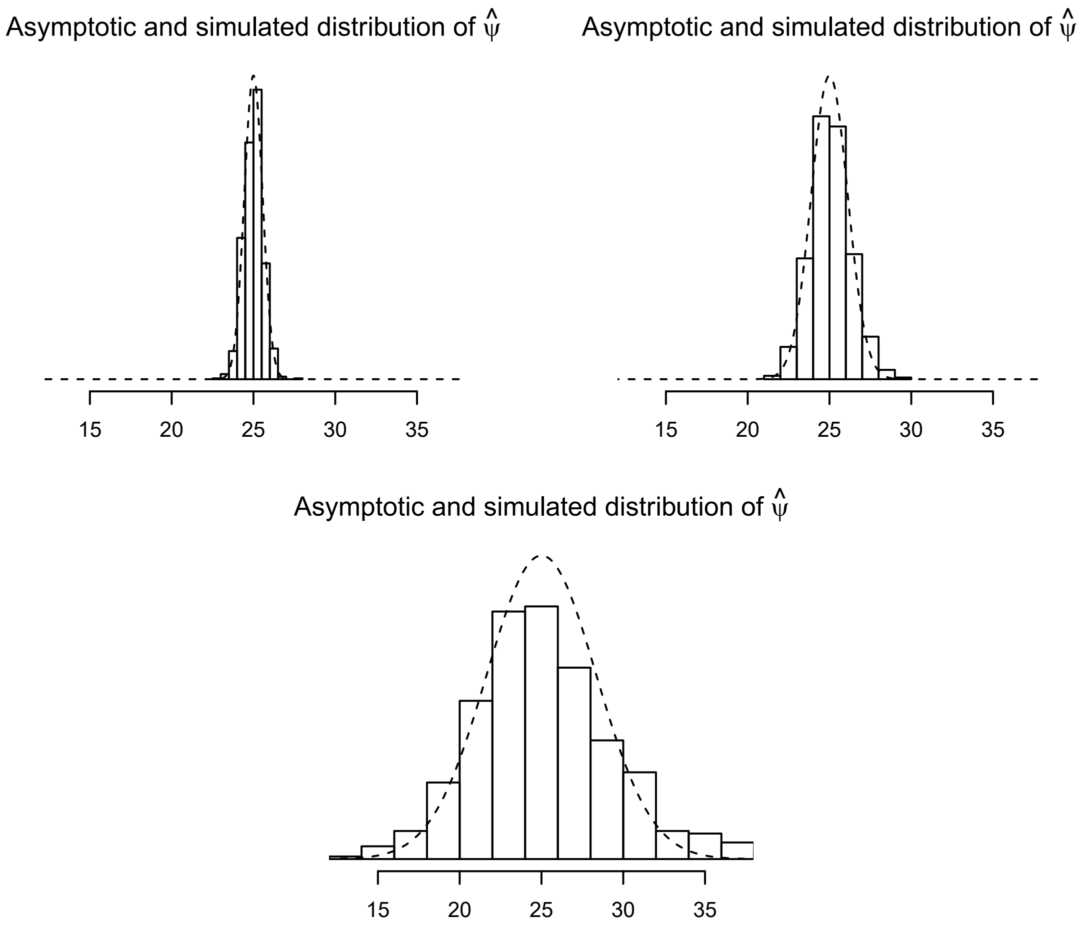

Figure 3.4: Asymptotic and simulated distribution of $\widehat{\psi}$ in linear PoSt model with unknown β_0 for $n = 50$, $\psi = 25$, $\boldsymbol{\beta} = (2, 2)^\top$ and $\sigma = 0.05$, 0.1, 0.3.

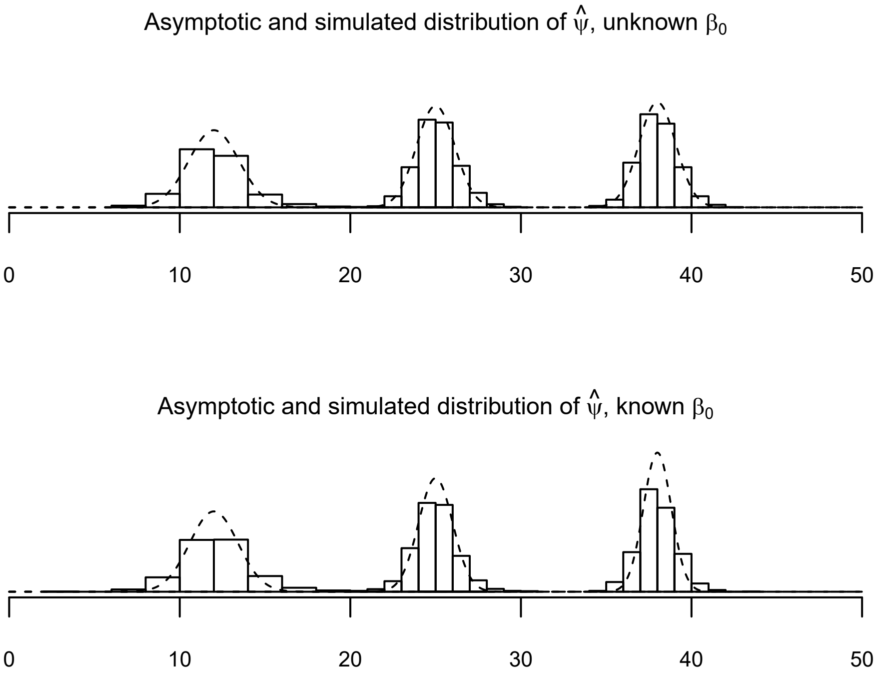

Figure 3.5: Asymptotic and simulated distribution of $\widehat{\psi}$ in linear PoSt model for different ψ, with known and unknown β_0, $n = 50$, $\psi = [n/4]$, $[n/2]$, $[3n/4]$, $\boldsymbol{\beta} = (0, 2)^\top$ and $\sigma = 0.1$.

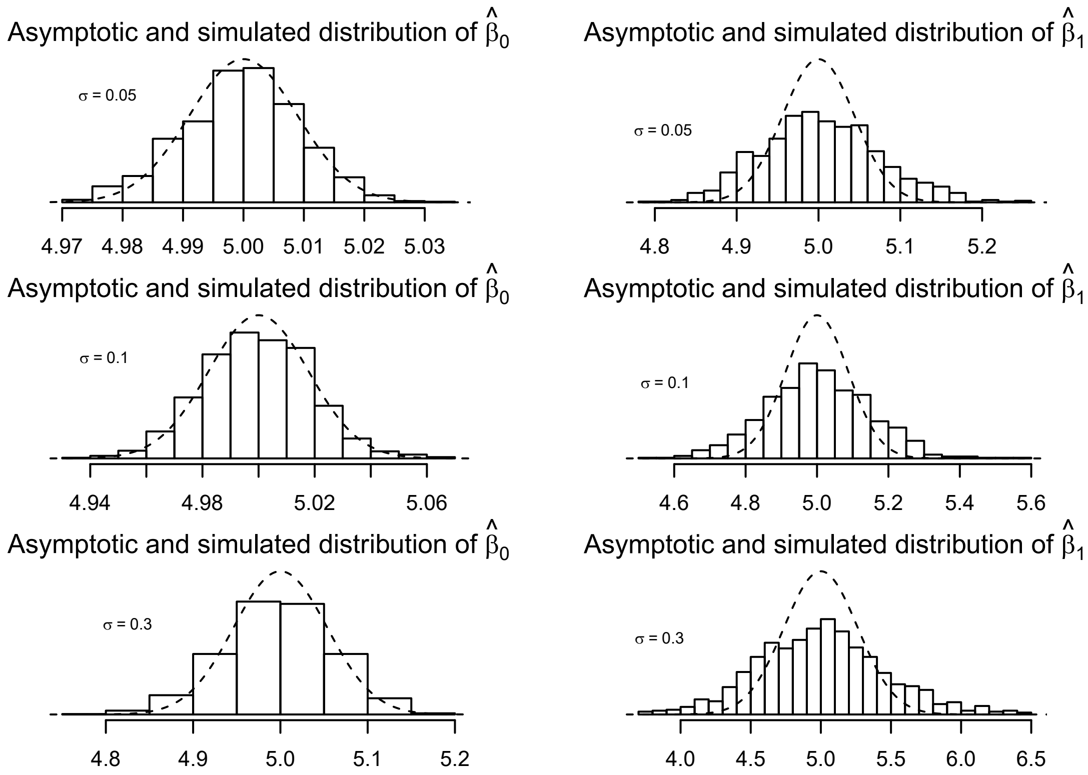

Figure 3.6: Asymptotic and simulated distribution of $\widehat{\boldsymbol{\beta}}$ in linear PoSt model with unknown β_0, $n = 50$, $\psi = 25$, $\boldsymbol{\beta} = (5, 5)^\top$ and $\sigma = 0.05$, 0.1, 0.3.

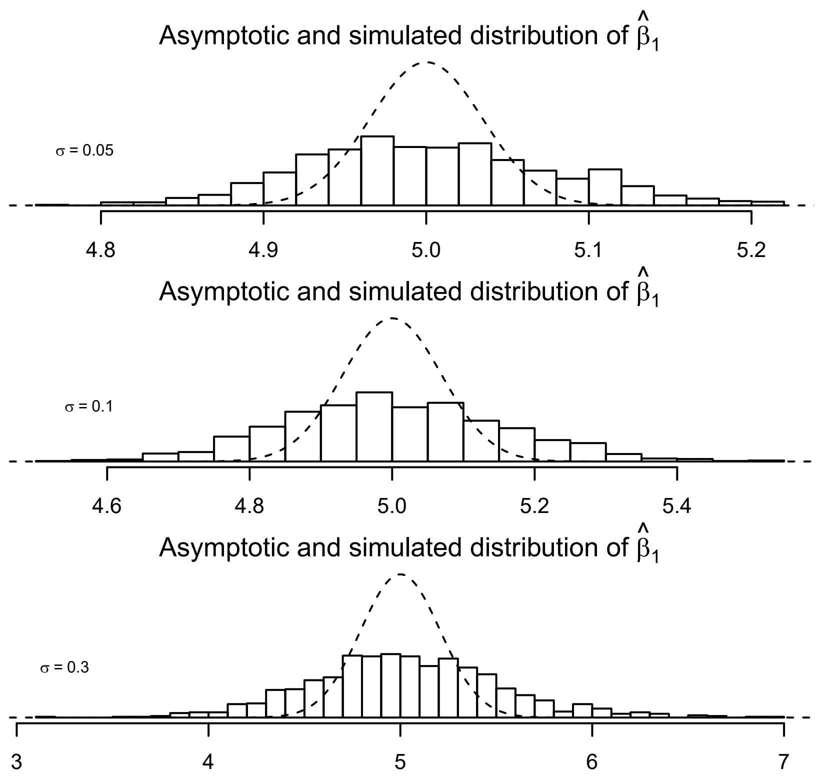

Figure 3.7: Asymptotic and simulated distribution of $\widehat{\beta_1}$ in linear PoSt model with known β_0, $n = 50$, $\psi = 25$, $\beta_1 = 5$ and $\sigma = 0.05$, 0.1, 0.3.

3.5.1 Confidence intervals coverage

Now, we run simulations to get the coverage of two-sided and right-sided asymptotic confidence interval for ψ given by (3.7) and (3.5). Further, we construct the intervals using bootstrap. The coverage and the average length of the interval are compared for the two methods.

We run the simulations in more setups, for all we set parameters to values $\beta_0 = 2$, $\beta_1 = 2$, $\sigma = 0.03$. Confidence intervals (CI) are all computed with $\alpha = 0.05$. The number of repetitions N and the number of bootstrap samples B are set to 1000. We try different sample sizes $n = 25, 50, 100$ and for each n we set change-point to $\psi = [n/4]$, $[n/2]$, $[3n/4]$. We generate the data $Z_1, \ldots, Z_n$ from model (3.1) with normally distributed errors e_i having $\operatorname{var} e_i = \sigma^2$. The bootstrap confidence interval for ψ is constructed as follows.

1. Estimate unknown parameters by $\widehat{\psi}$, $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\sigma}$ using data $Z_1, \ldots, Z_n$.
2. For $b = 1, \ldots, B$, where B is the number of bootstrap samples:
   (a) Sample $e^*_{1,b}, \ldots, e^*_{n,b}$ from $\mathrm{N}(0, \widehat{\sigma}^2)$.
   (b) Calculate $Z^*_{i,b} = \widehat{\beta}_0 + \widehat{\beta}_1 \left(\frac{\widehat{\psi}-i}{n}\right)^+ + e^*_{i,b}$ for $i = 1, \ldots, n$.
   (c) Estimate ψ from $Z^*_{1,b}, \ldots, Z^*_{n,b}$ and denote it $\widehat{\psi}^*_b$.
3. Let $q^*_{n,B}(\alpha)$ denote the α sample quantile calculated from $\widehat{\psi}^*_1, \ldots, \widehat{\psi}^*_B$. Construct bootstrap confidence interval for ψ as

$$\left(2\,\widehat{\psi} - q^*_{n,B}(1-\alpha/2),\ 2\,\widehat{\psi} - q^*_{n,B}(\alpha/2)\right).$$

By bootstrap, we want to approximate the quantiles of $\sqrt{n}\left(\widehat{\psi} - \psi\right)$ by using quantiles of $\sqrt{n}\left(\widehat{\psi}^* - \widehat{\psi}\right)$, denoted by $r^*_n(\alpha)$. These quantiles can be further approximated by $r^*_{n,B}(\alpha)$ calculated from B Monte Carlo simulations. Then, the $(1-\alpha)\,100\%$ bootstrap confidence interval for ψ has form

$$\left(\widehat{\psi} - \frac{r^*_{n,B}\left(1-\alpha/2\right)}{\sqrt{n}}, \widehat{\psi} - \frac{r^*_{n,B}\left(\alpha/2\right)}{\sqrt{n}}\right).$$

Denoting $q^*_{n,B}(\alpha)$ the $\alpha-$quantile calculated from $\widehat{\psi}^*_1, \ldots, \widehat{\psi}^*_B$ we then obtain $r^*_{n,B}\left(\alpha\right) = \sqrt{n}\left(q^*_{n,B}(\alpha) - \widehat{\psi}\right)$ and we can rewrite the interval as

$$\left(2\,\widehat{\psi} - q^*_{n,B}(1-\alpha/2),\ 2\,\widehat{\psi} - q^*_{n,B}(\alpha/2)\right).$$

Results for the two-sided confidence interval are summarised in Table 3.1. For lowest $n = 25$, the coverage of both types of CI is lower than 95% and it is increasing with increasing ψ. Bootstrap CI have higher coverage by 0.1% than asymptotic CI for $\psi = 6, 12$ and lower coverage for $\psi = 19$. The bootstrap CI is also slightly wider. For $n = 50, 100$, the coverages are slightly smaller than 95%. For both n, the coverage is similar for both types of CI. The average length is similar, again slightly larger for bootstrap CI and it decreases with increasing ψ.

In general, bootstrap CI have better coverage for $\psi = [n/4], [n/2]$ for all chosen n but all the results are very similar.

The results for the right-sided interval (3.5) are summarised in Table 3.2, the setup is the same as for two-sided interval. Here, we monitor mean distance $c_U - \psi$. For the asymptotic interval, this distance is equal to the average of $c_{Ui} - \psi$ where c_{Ui} denotes the lower bound of the confidence interval interval for the i-th repetition, $i = 1, \ldots, N$. For the bootstrap interval, it is equal to the average of $c_{Ui} - \psi$ across all repetitions where for each repetition $i = 1, \ldots, N$, c_{Ui} denotes the upper bound of bootstrap CI calculated from B bootstrap samples. In general, the coverages are similar to coverages for two-sided CI.

n	ψ	**Coverage asymptotic CI** [%]	**Coverage bootstrap CI** [%]	**Mean length asymptotic CI**	**Mean length bootstrap CI**
	6	88.8	90.9	0.77	0.86
$n = 25$	12	90.8	90.9	0.58	0.61
	19	93.3	93.0	0.56	0.57
	12	92.2	94.0	1.13	1.22
$n = 50$	25	94.1	94.4	0.85	0.86
	38	93.2	93.0	0.82	0.83
	25	94.9	95.0	1.61	1.66
$n = 100$	50	94.3	95.2	1.22	1.23
	75	94.5	94.7	1.18	1.18

Table 3.1: Coverage and average length of a two-sided confidence interval (CI) for ψ in linear PoSt model based on asymptotic distribution and bootstrap for $\boldsymbol{\beta} = (2, 2)^\top$, $\sigma = 0.02$, $B = 1000$.

n	ψ	**Coverage asymptotic CI** [%]	**Coverage bootstrap CI** [%]	**Mean distance** $c_U - \psi$ **asymptotic**	**Mean distance** $c_U - \psi$ **bootstrap**
	6	88.0	88.7	0.30	0.33
$n = 25$	12	89.7	90.2	0.24	0.25
	19	93.5	93.0	0.24	0.24
	12	92.6	92.1	0.47	0.49
$n = 50$	25	94.5	94.3	0.36	0.37
	38	93.7	93.6	0.34	0.34
	25	94.6	95.6	0.68	0.70
$n = 100$	50	94.2	93.8	0.52	0.52
	75	94.3	93.8	0.50	0.50

Table 3.2: Coverage of a right-sided confidence interval (CI) $(-\infty, c_U)$ for ψ and average distance $c_U - \psi$ in linear PoSt model based on asymptotic distribution and bootstrap for $\boldsymbol{\beta} = (2, 2)^\top$, $\sigma = 0.02$, $B = 1000$.

4. Heteroscedasticity in linear PoSt model

In applications, the assumption of homoscedasticity is not always met. We often face a situation when the variance of random error varies accross time $i = 1, \dots, n$. For example, observations may have larger variance when the trends occurs and the variance may decrease as they stabilise. Therefore, it is convenient to collect multiple observations at each time i to verify the assumption of homoscedasticity or to adjust the methods if heteroscedasticity may be present.

We work with the PoSt model

$$Z_i = \beta_0 + \beta_1 \left(\frac{\psi - i}{n}\right)^+ + e_i \tag{4.1}$$

from Chapter 3 but the described procedure to derive the estimates and bootstrap under is analogous for a model with general polynomial trend from Chapter 2. Model with repeated measurements can be written as

$$Z_{ij} = \beta_0 + \beta_1 \left(\frac{\psi - i}{n}\right)^+ + e_{ij}, \quad i = 1, \dots, n \text{ and } j = 1, \dots, m_i. \tag{4.2}$$

Here we assume having different number of observations m_i at each time i. Under the homoscedasticity assumption it holds $\operatorname{var} e_{ij} = \sigma^2$, under heteroscedasticity assumption we have $\operatorname{var} e_{ij} = \sigma_i^2$.

Under homoscedasticity and if $m_i = m$ for all i, we can proceed similarly as in previous chapters by taking sample means of observations at each i instead of individual observations. For each i the sample mean has variance σ^2/m.

Now, let us assume heteroscedasticity and possibly different numbers of observations m_i at each time i. Under this setup, we work with sample means $\overline{Z_i} = \sum_{j=1}^{m_i} Z_{ij}/m_i$ at each time $i = 1, \dots, n$.

We solve

$$\min_{\substack{\beta_0, \beta_1 \in \mathbb{R} \\ p \in (1,n)}} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \left(Z_{ij} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{1}{\sigma_i^2}$$

and we can rewrite the minimised expression as

$$\sum_{i=1}^{n} \sum_{j=1}^{m_i} \left(Z_{ij} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{1}{\sigma_i^2} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \left(Z_{ij} - \overline{Z_i} + \overline{Z_i} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{1}{\sigma_i^2} =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{\left(Z_{ij} - \overline{Z_i}\right)^2}{\sigma_i^2} + \sum_{i=1}^{n} \left(\overline{Z_i} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{m_i}{\sigma_i^2}.$$

Since the first part does not depend on p and β_0, β_1 we can rewrite the minimisation task as

$$\min_{\substack{\beta_0, \beta_1 \in \mathbb{R} \\ p \in (1,n)}} \sum_{i=1}^{n} \left(\overline{Z_i} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{1}{w_i} \tag{4.3}$$

denoting $w_i = \frac{\sigma_i^2}{m_i} = \operatorname{var} \overline{Z_i}$.

Denoting $\overline{\boldsymbol{Z}} = \left(\overline{Z_1}, \ldots, \overline{Z_n}\right)^\top$, the minimised term can be written in matrix form as

$$\sum_{i=1}^{n} \left(\overline{Z_i} - \beta_0 - \beta_1 x_{p,i}\right)^2 \frac{1}{w_i} = \left(\overline{\boldsymbol{Z}} - \mathbb{X}_{p\cdot}\boldsymbol{\beta}\right)^\top \mathbb{W} \left(\overline{\boldsymbol{Z}} - \mathbb{X}_{p\cdot}\boldsymbol{\beta}\right)$$

where

$$\mathbb{W} = \begin{pmatrix} \frac{1}{w_i} & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & \frac{1}{w_n} \end{pmatrix} = \begin{pmatrix} \frac{m_1}{\sigma_1^2} & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & \frac{m_n}{\sigma_n^2} \end{pmatrix}.$$

Matrix $\mathbb{W}^{-1}$ is the variance matrix of vector $\overline{\boldsymbol{Z}}$.

As the variances σ_i^2 are unknown, we have to estimate them first by sample variances $\widehat{\sigma}_i^2 = \frac{1}{m_i - 1}\sum_{j=1}^{m_i}\left(Z_{ij} - \overline{Z_i}\right)^2$. These estimators are used to estimate w_i by $\widehat{w_i} = \widehat{\sigma}_i^2/m_i$ and matrix $\mathbb{W}$ by matrix $\widehat{\mathbb{W}}$ consisting of estimators $\widehat{w_i}$.

Therefore instead of (4.3) we solve

$$\min_{\substack{\beta_0,\beta_1 \in \mathbb{R} \\ p \in (1,n)}} \quad \left(\overline{\boldsymbol{Z}} - \mathbb{X}_{p\cdot}\boldsymbol{\beta}\right)^\top \widehat{\mathbb{W}} \left(\overline{\boldsymbol{Z}} - \mathbb{X}_{p\cdot}\boldsymbol{\beta}\right). \tag{4.4}$$

Proceeding as in Chapter 2, we obtain

$$\begin{aligned} \widehat{\psi} &= \underset{p \in (1,n)}{\arg\min}\; \overline{\boldsymbol{Z}}^\top \widehat{\mathbb{W}} \overline{\boldsymbol{Z}} - \overline{\boldsymbol{Z}}^\top \widehat{\mathbb{W}} \mathbb{X}_{p\cdot} \left(\mathbb{X}_{p\cdot}^\top \widehat{\mathbb{W}} \mathbb{X}_{p\cdot}\right)^{-1} \mathbb{X}_{p\cdot}^\top \widehat{\mathbb{W}} \overline{\boldsymbol{Z}} \\ &= \underset{p \in (1,n)}{\arg\max}\; \overline{\boldsymbol{Z}}^\top \widehat{\mathbb{W}} \mathbb{X}_{p\cdot} \left(\mathbb{X}_{p\cdot}^\top \widehat{\mathbb{W}} \mathbb{X}_{p\cdot}\right)^{-1} \mathbb{X}_{p\cdot}^\top \widehat{\mathbb{W}} \overline{\boldsymbol{Z}} \end{aligned} \tag{4.5}$$

For the estimator of $\boldsymbol{\beta}$ we have

$$\widehat{\boldsymbol{\beta}} = \left(\mathbb{X}_{\widehat{\psi}\cdot}^\top \widehat{\mathbb{W}} \mathbb{X}_{\widehat{\psi}\cdot}\right)^{-1} \mathbb{X}_{\widehat{\psi}\cdot}^\top \widehat{\mathbb{W}} \overline{\boldsymbol{Z}}.$$

For the linear case, we can rewrite the formula for $\widehat{\psi}$ and we obtain

$$\widehat{\psi} = \underset{p \in (1,n)}{\arg\max} \frac{\left(\sum_{i=1}^{n} \overline{Z_i}/\widehat{w_i}\left(x_{p,i} - \frac{\sum_{i=1}^{n} x_{p,i}/\widehat{w_i}}{\sum_{i=1}^{n} 1/\widehat{w_i}}\right)\right)^2}{\sum_{i=1}^{n} x_{p,i}^2/\widehat{w_i} - \frac{\left(\sum_{i=1}^{n} x_{p,i}/\widehat{w_i}\right)}{\sum_{i=1}^{n} 1/\widehat{w_i}}}.$$

In Figure 4.1, variability of the simulated data varying at each time is visualised using boxplots calculated from repeated observations at each time i. The variability differs particularly for small i and gets smaller as the trend diminishes. The parameters are estimated by $\widehat{\psi}$ and $\widehat{\boldsymbol{\beta}}$.

For heteroscedasticity and varying numbers of observations m_i, we use bootstrap to approximate the distribution of the estimated parameters and to construct confidence intervals.

The bootstrap procedure runs as follows:

1. Given data $Z_{i,j}, i = 1, \ldots, n, j = 1, \ldots, m_i$, estimate $\widehat{\sigma}_i^2$, $\widehat{\mathbb{W}}$, $\widehat{\psi}$ and $\widehat{\beta}$ using formulae above.

2. For $b = 1, \dots, B$, where B is the number of bootstrap samples:

   (a) Sample $e^*_{1,b}, \dots, e^*_{n,b}$ where $e^*_{i,b}$ is sampled from $\mathrm{N}(0,\, \widehat{\sigma}^2_i/m_i)$.

   (b) Calculate $\overline{Z}^*_{i,b} = \widehat{\beta}_0 + \widehat{\beta}_1 \left(\frac{\widehat{\psi}-i}{n}\right)^+ + e^*_{i,b}$ for $i = 1, \dots, n$.

   (c) Estimate ψ by $\widehat{\psi}^*_b$ calculated from $\overline{Z}^*_{1,b}, \dots, \overline{Z}^*_{n,b}$ as above using $\widehat{\mathbb{W}}$.

3. Let $q^*_{n,B}(\alpha)$ denote the α sample quantile calculated from $\widehat{\psi}^*_1, \dots, \widehat{\psi}^*_B$. Construct bootstrap confidence interval for ψ as

$$\left(2\,\widehat{\psi} - q^*_{n,B}(1-\alpha/2),\ 2\,\widehat{\psi} - q^*_{n,B}(\alpha/2)\right).$$

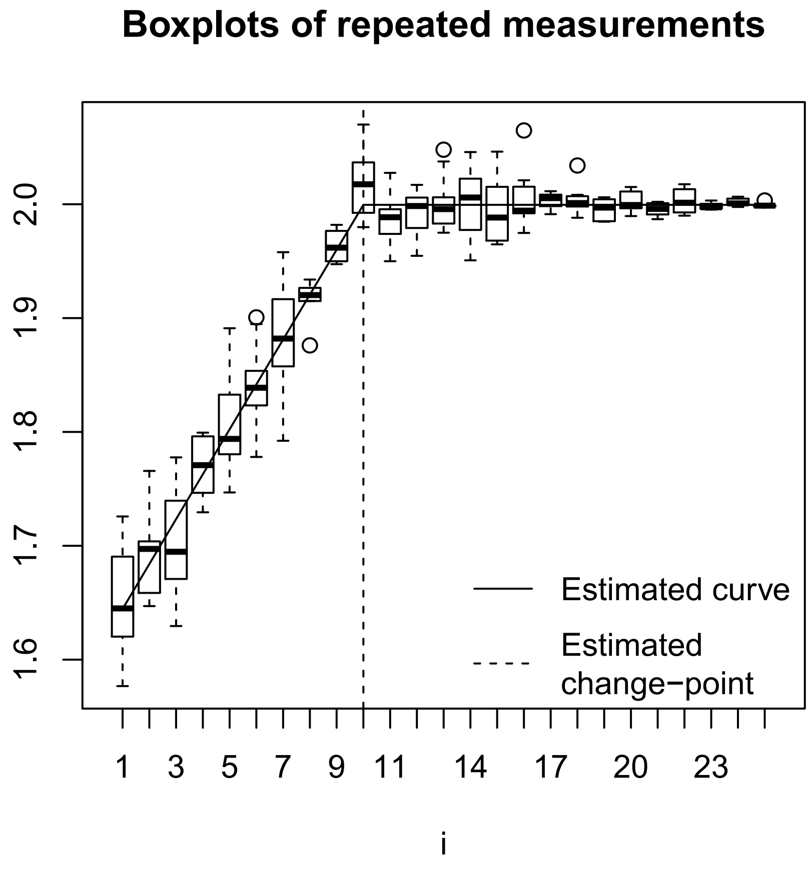

Figure 4.1: Boxplots of repeated measurements at each time i from heteroscedastic model (4.2) with estimated change-point ψ and estimated curve using $\widehat{\boldsymbol{\beta}}$.

5. Nonlinear PoSt models

In this chapter, we extend linear *PoSt* model to nonlinear PoSt models using results about quadratic and E_{max} model from Section 2.2 and 2.3. Then, we run simulations to visualise the asymptotic distribution of estimators together with histograms of simulated values for different parameter choices. Coverage of confidence intervals based on asymptotic results and/or bootstrap is also calculated and selected methods are compared.

5.1 Quadratic model

Quadratic change-point model discussed in Section 2.2 is already formulated in *PoSt* context, i.e. assuming trend up to change-point ψ and stable data after ψ. Therefore, we can use results from theorems in Section 2.2.

Assuming β_0 is unknown and $\beta_1 \neq 0$ the asymptotic both-sided $(1-\alpha)\,100\%$ confidence interval for the change-point ψ from (2.15) has form

$$\left(\widehat{\psi} - u_{1-\frac{\alpha}{2}} \frac{\sigma\sqrt{n}}{\widehat{\beta}_1} \sqrt{\frac{9-5\theta_\psi}{\theta_\psi\left(1-\theta_\psi\right)}},\ \widehat{\psi} + u_{1-\frac{\alpha}{2}} \frac{\sigma\sqrt{n}}{\widehat{\beta}_1} \sqrt{\frac{9-5\theta_\psi}{\theta_\psi\left(1-\theta_\psi\right)}}\right) \tag{5.1}$$

and the right-sided confidence interval has form

$$\left(-\infty,\ \widehat{\psi} + u_{1-\alpha} \frac{\sigma\sqrt{n}}{\widehat{\beta}_1} \sqrt{\frac{9-5\theta_\psi}{\theta_\psi\left(1-\theta_\psi\right)}}\right) \tag{5.2}$$

As for linear *PoSt* model in Chapter 3 we first visualise the asymptotic and simulated distribution of the parameter estimators and then calculate the coverage of the confidence intervals based on the asymptotic results and bootstrap.

Similarly as in linear *PoSt* model, for small values of σ, the simulated values of $\widehat{\psi}$ are not so dispersed as for larger σ, see Figure 5.1 created with known $\beta_0 = 0$ and $n = 50$, $\psi = 25$, $\beta_1 = \beta_2 = 5$ and $\sigma = 0.05, 0.1, 0.3$. In all cases, the simulated values provide a reasonable approximation of the asymptotic distribution.

As in linear *PoSt* model, simulated values are most dispersed for the lowest ψ representing the earliest location of the change-points, see Figure 5.2 calculated with $\boldsymbol{\beta} = (0,5,5)^\top$, $n = 50$ and $\sigma = 0.1$. For larger ψ, the simulated values are not so dispersed. The results are similar for both the cases with known and unknown β_0.

Next, we plot similar plots for estimators of β_1, β_2 in the case with known β_0 and $\beta_1 \neq 0$, see Figure 5.3. For the smallest chosen σ the asymptotic and the simulated distribution are very similar. The asymptotic distributions of β_2 have a slightly larger variance then variance of β_1 and the simulated values are more dispersed then for β_1. For the highest chosen σ, some of the simulated $\widehat{\beta_2}$ underestimate the real β_2, few of them are even negative. We also get some large values of $\widehat{\beta_1}$ with the same σ.

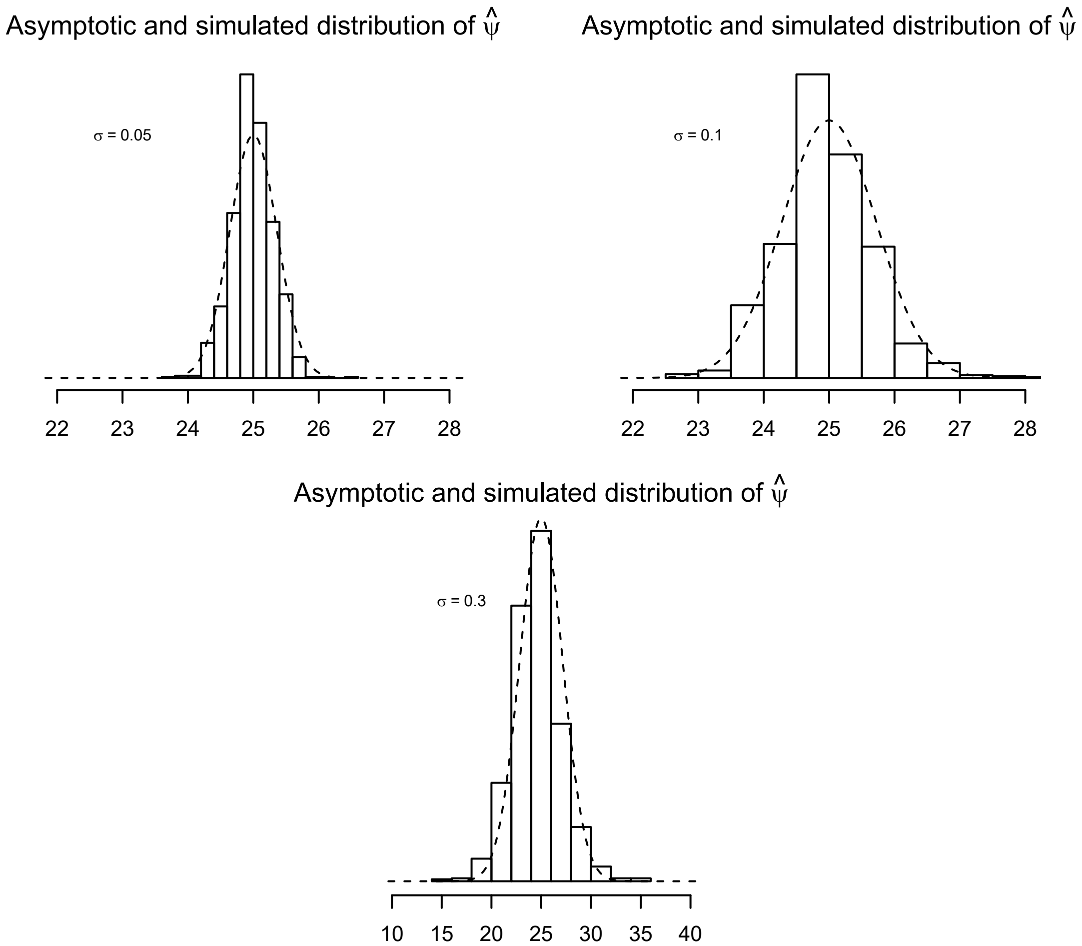

Figure 5.1: Asymptotic and simulated distribution of $\widehat{\psi}$ in quadratic PoSt model with known $\beta_0 = 0$ for $n = 50, \psi = 25, \beta_1 = \beta_2 = 5$ and $\sigma = 0.05,\ 0.1,\ 0.3$.

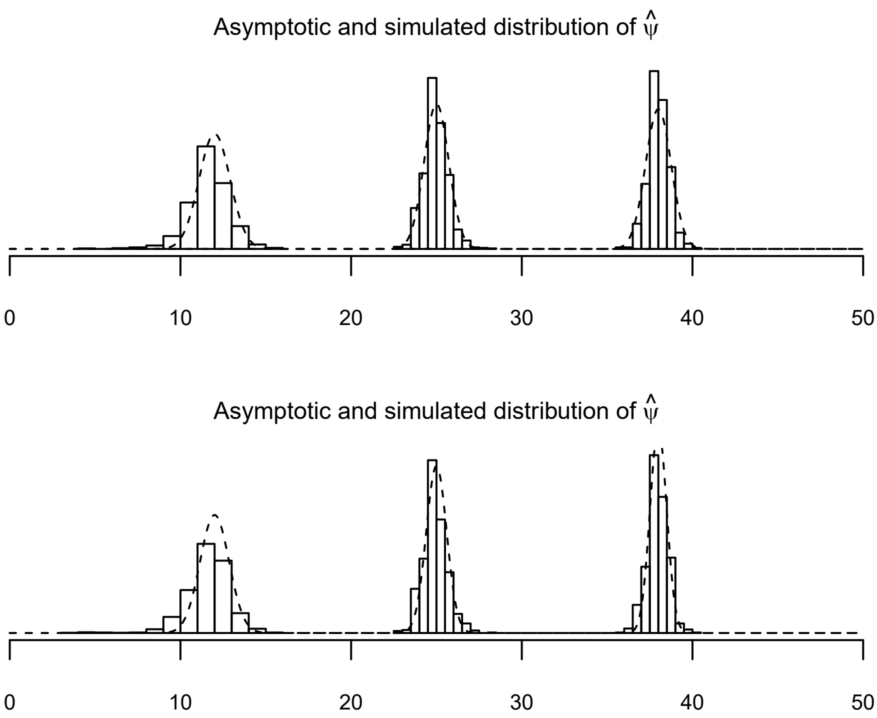

Figure 5.2: Asymptotic and simulated distribution of $\widehat{\psi}$ in quadratic PoSt model for different ψ, with known and unknown β_0, $n = 50$, $\psi = [n/4], [n/2], [3n/4]$, $\boldsymbol{\beta} = (0, 5, 5)^{\top}$ and $\sigma = 0.1$.

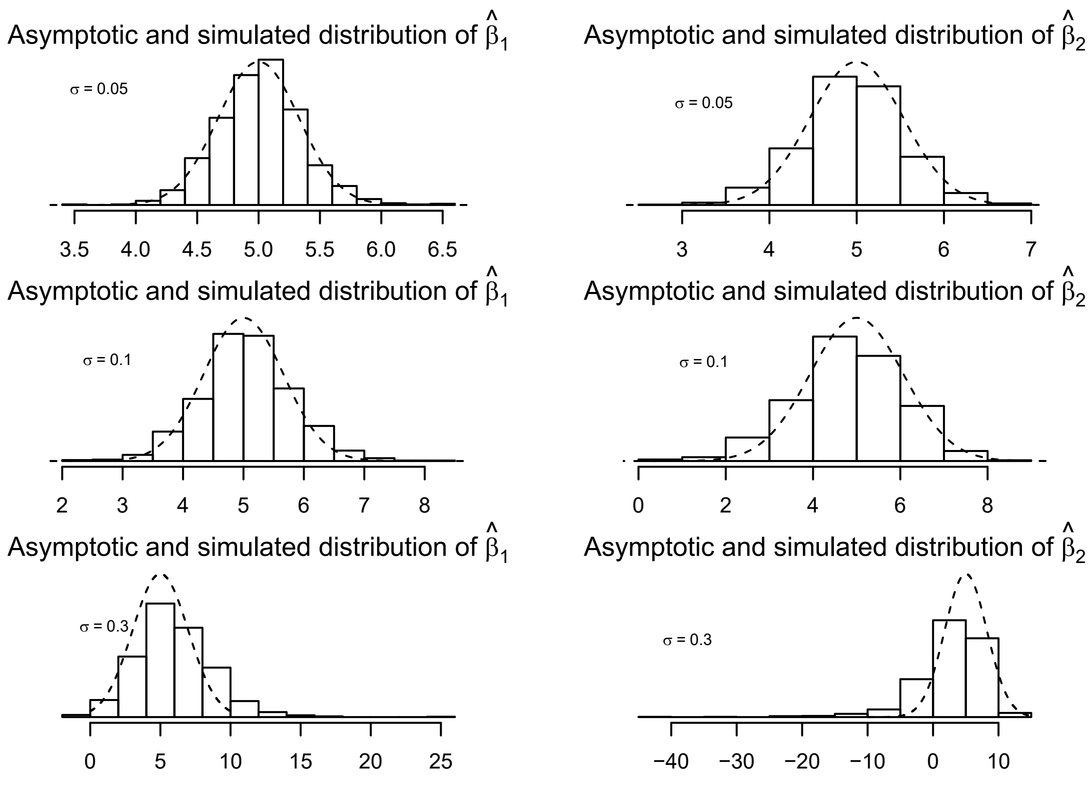

Figure 5.3: Asymptotic and simulated distribution of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ in quadratic PoSt model with known β_0, $n = 50$, $\psi = 25$, $\beta_1 = \beta_2 = 5$ and $\sigma = 0.05$, 0.1, 0.3.

5.1.1 Confidence intervals coverage

Now, we run simulations to calculate the coverage of the two-sided and right-sided confidence interval for change-point ψ in based on asymptotic distribution and bootstrap. We set $n = 25, 50, 100$ and for every n we have $\psi = [n/4], [n/2], [3n/4]$. For all combinations of n and ψ, we assume β_0 to be unknown and we set $\beta_0 = \beta_1 = \beta_2 = 3$, $\sigma = 0.03$, number of repetitions $N = 1000$ and number of bootstrap samples $B = 700$. We set smaller B than in simulations at linear *PoSt* model because estimation of the change-point in quadratic model is more time consuming. For the linear model, we estimated ψ as the argument of the maximum of (3.2) using Brent minimisation method (with negative sign) implemented in R, see Brent [2013]. For quadratic model, results of this method were poor therefore we chose conjugate gradients method (see Fletcher and Reeves [1964]) which appeared to be more time consuming but estimators $\widehat{\psi}$ were more accurate.

The bootstrap procedure to construct the two-sided confidence interval for ψ runs as follows.

1. Estimate unknown parameters by $\widehat{\psi}$, $\widehat{\beta}_0$, $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $\widehat{\sigma}$,
2. For $b = 1, \ldots, B$, where B is the number of bootstrap samples
   (a) Sample $e^*_{1,b}, \ldots, e^*_{n,b}$ from $\mathrm{N}(0, \widehat{\sigma}^2)$
   (b) Calculate $Z^*_{i,b} = \widehat{\beta}_0 + \widehat{\beta}_1 \, x_{\widehat{\psi},i} + \widehat{\beta}_2 \, x^2_{\widehat{\psi},i} + e^*_{i,b}$ for $i = 1, \ldots, n$
   (c) Estimate ψ from $Z^*_{1,b}, \ldots, Z^*_{n,b}$ and denote it $\widehat{\psi}^*_b$
3. Let $q^*_{n,B}(\alpha)$ denote the α sample quantile calculated from $\widehat{\psi}^*_1, \ldots, \widehat{\psi}^*_B$. Construct bootstrap confidence interval for ψ as

$$\left(2\,\widehat{\psi} - q^*_{n,B}(1-\alpha/2),\ 2\,\widehat{\psi} - q^*_{n,B}(\alpha/2)\right).$$

Results are summarised in Table 5.1. The two-sided confidence interval (5.1) based on the asymptotic distribution has the lowest coverage for the lowest ψ at each n and the coverage increases as ψ increases. For $\psi = [3n/4]$, the coverage is higher than 95%. For the lowest $\psi = [n/4]$ we have the highest coverage for $n = 50$.

Confidence interval based on bootstrap has lower coverage than asymptotic confidence intervals for $\psi = [n/2]$, $[3n/4]$ and higher coverage than asymptotic confidence interval for $\psi = [n/4]$ except for $n = 50$. Similarly as for the asymptotic CI, the coverage is lowest for the lowest $\psi = [n/4]$. Also, the bootstrap interval is wider for $\psi = [n/4]$ than the asymptotic interval and shorter for the other location of ψ.

We calculate the coverage also for the right-sided interval (5.2) based on asymptotic results and right-sided interval based on bootstrap. Results are summarised in Table 5.2, the coverage is similar as for the two-sided interval. The coverage is lowest for the lowest $\psi = [n/4]$ and it increases with increasing ψ. The bootstrap has better coverage for $\psi = [n/4]$ than the asymptotic confidence interval and slightly lower for other selected values of ψ.

n	ψ	**Coverage asymptotic CI** [%]	**Coverage bootstrap CI** [%]	**Mean length asymptotic CI**	**Mean length bootstrap CI**
$n = 25$	6	77.1	88.3	1.11	1.87
	12	92.8	90.7	0.91	0.87
	19	97.2	90.9	0.95	0.74
$n = 50$	12	89.7	88.3	1.68	1.69
	25	95.4	94.2	1.35	1.24
	38	98.1	95.4	1.41	1.08
$n = 100$	25	81.2	83.8	2.26	1.82
	50	95.5	91.6	1.93	1.63
	75	98.5	93.4	2.02	1.51

Table 5.1: Coverage and average length of a two-sided confidence interval (CI) for ψ in quadratic model based on asymptotic distribution and bootstrap for $\beta_0 = \beta_1 = \beta_2 = 3$, $\sigma = 0.03$, $B = 700$.

n	ψ	**Coverage asymptotic CI** [%]	**Coverage bootstrap CI** [%]	**Mean distance** $c_U - \psi$ **asymptotic**	**Mean distance** $c_U - \psi$ **bootstrap**
$n = 25$	6	72.1	88.1	0.24	0.75
	12	90.1	87.7	0.35	0.34
	19	95.5	91.0	0.39	0.31
$n = 50$	12	86.6	89.2	0.57	0.69
	25	94.7	93.5	0.55	0.49
	38	97.6	93.7	0.57	0.42
$n = 100$	25	73.5	82.5	0.35	0.60
	50	93.0	91.9	0.68	0.68
	75	97.0	93.2	0.79	0.64

Table 5.2: Coverage of a right-sided confidence interval (CI) for ψ and average distance $c_U - \psi$ in quadratic model based on asymptotic distribution and bootstrap for $\beta_0 = \beta_1 = \beta_2 = 3$, $\sigma = 0.03$, $B = 700$.

5.2 E_{max} model

In this section, we examine the E_{max} model with change-point ψ discussed in Section 2.3 which can be used also in *PoSt* context. The model

$$Z_i = \beta_{lowAs} + \beta_{incr} \frac{1}{1 + \frac{\gamma}{\min(i,\psi)}} + e_i$$

$$= \beta_{lowAs} + \beta_{incr} \ + x_{\psi,i;\gamma} + e_i$$

represents non-polynomial generalisation of the *PoSt* model. For the E_{max} model we do not have asymptotic results, therefore we visualise simulated distributions together with the real value of parameters.

As discussed in Section 2.3 in Chapter 2, the parameter β_{lowAs} represents the lower asymptote i.e. the situation with no effect. We often want to estimate the mean value of the response after the change-point, which was denoted by β_0 in polynomial *PoSt* models. In E_{max} model with change-point, this value can be calculated as $\beta_{lowAs} + \beta_{incr}/\left(1 + \frac{\gamma}{\psi}\right)$ and it equals $\mathsf{E}\ Z_\psi$. When generating data from E_{max} model (e.g. for the figures or when constructing the confidence intervals), one of inputs is the value β_0, from which we calculate β_{lowAs} such that, after the change-point, the mean value of the response will be the given β_0. The data are then generated with this calculated β_{lowAs} and estimation proceeds as described in Section 2.3.

We plot the simulated distribution for different combinations of parameters, similarly as in previous section. For $\sigma = 0.05$, the simulated distribution of $\widehat{\psi}$ is centered in the true ψ and it estimates ψ nicely, see Figure 5.4 created with $n = 50$, $\psi = 25$, $\beta_{incr} = 2$, $\gamma = 25$ and $\sigma = 0.05, 0.1, 0.3$. For higher values of σ, the estimates of the change-points are more dispersed.

In E_{max} model, we observe different behaviour with different location of change-point compared to linear and quadratic *PoSt* model. We have the smallest variance of the simulated distribution for the lowest ψ and the most dispersed values of $\widehat{\psi}$ for the highest ψ, see Figure 5.5 created with $n = 50$, $\psi = [n/4]$, $[n/2]$, $[3n/4]$, $\beta_{incr} = 2$, $\gamma = 25$ and $\sigma = 0.1$.

Next, we look at the simulated distribution of $\widehat{\beta}_{lowAs}$ and $\widehat{\beta}_{incr}$, see Figure 5.6 created with $n = 50$, $\psi = 25$, $\beta_{incr} = 2$, $\gamma = [n/2]$ and $\sigma = 0.05, 0.1, 0.3$.. The parameter β_{lowAs} was calculated such that the mean value after the change-point equals 5, as discussed above. The simulated distribution of $\widehat{\beta}_{lowAs}$ is centered around the true value and it is more dispersed as σ increases. For $\widehat{\beta}_{incr}$, the situation is similar, but the simulated histogram seems slightly nonsymmetric, we have more values exceeding the true β_{incr} than we would expect.

5.2.1 Confidence intervals coverage

The confidence intervals for E_{max} model can be constructed using bootstrap. We describe the bootstrap process to construct the two-sided confidence interval for the change-point ψ.

1. Estimate $\widehat{\psi}$, $\widehat{\gamma}$, $\widehat{\boldsymbol{\beta}} = (\beta_{lowAs}, \beta_{incr})^\top$ and $\widehat{\sigma}^2$ from data $Z_1, \ldots, Z_n$.
2. For $b = 1, \ldots, B$ where B is selected number of bootstrap samples:

(a) Sample $e^*_{1,b}, \dots, e^*_{n,b}$ from $\mathrm{N}(0, \widehat{\sigma}^2)$.

(b) Calculate $Z^*_{i,b} = \widehat{\beta}_{lowAs} + \widehat{\beta}_{incr}\, x_{\widehat{\psi},i;\widehat{\gamma}} + e^*_{i,b}$ for $i = 1, \dots, n$.

(c) Estimate ψ from $Z^*_{1,b}, \dots, Z^*_{n,b}$ and denote it $\widehat{\psi}^*_b$.

3. Let $q^*_{n,B}(\alpha)$ denote the α sample quantile calculated from $\widehat{\psi}^*_1, \dots, \widehat{\psi}^*_B$. Construct bootstrap confidence interval for ψ as

$$\left(2\,\widehat{\psi} - q^*_{n,B}(1-\alpha/2),\ 2\,\widehat{\psi} - q^*_{n,B}(\alpha/2)\right).$$

In Figure 5.7 we see generated data with $n = 50$, $\psi = 25$, $\gamma = 25$, $\beta_{incr} = 5$, $\sigma = 0.1$ and $\beta_{lowAs} = 0.2$ with 95% right-sided confidence interval for ψ constructed using bootstrap with $B = 1000$.

We run simulations to calculate the coverage of the two-sided and right-sided confidence interval for $n = 25$, 50, 100, $\psi = [n/4]$, $[n/2]$, $[3n/4]$. We further set $B = 1000$, $N = 1000$, $\sigma = 0.03$, $\gamma = \psi$, $\beta_{incr} = 3$ and we calculate β_{lowAs} such that the mean value of the response after the change-points equals 3.

Results for the two-sided confidence interval are summarised in Table 5.3. In general, the coverage in almost all cases is below 95%. For $n = 25$, the coverage

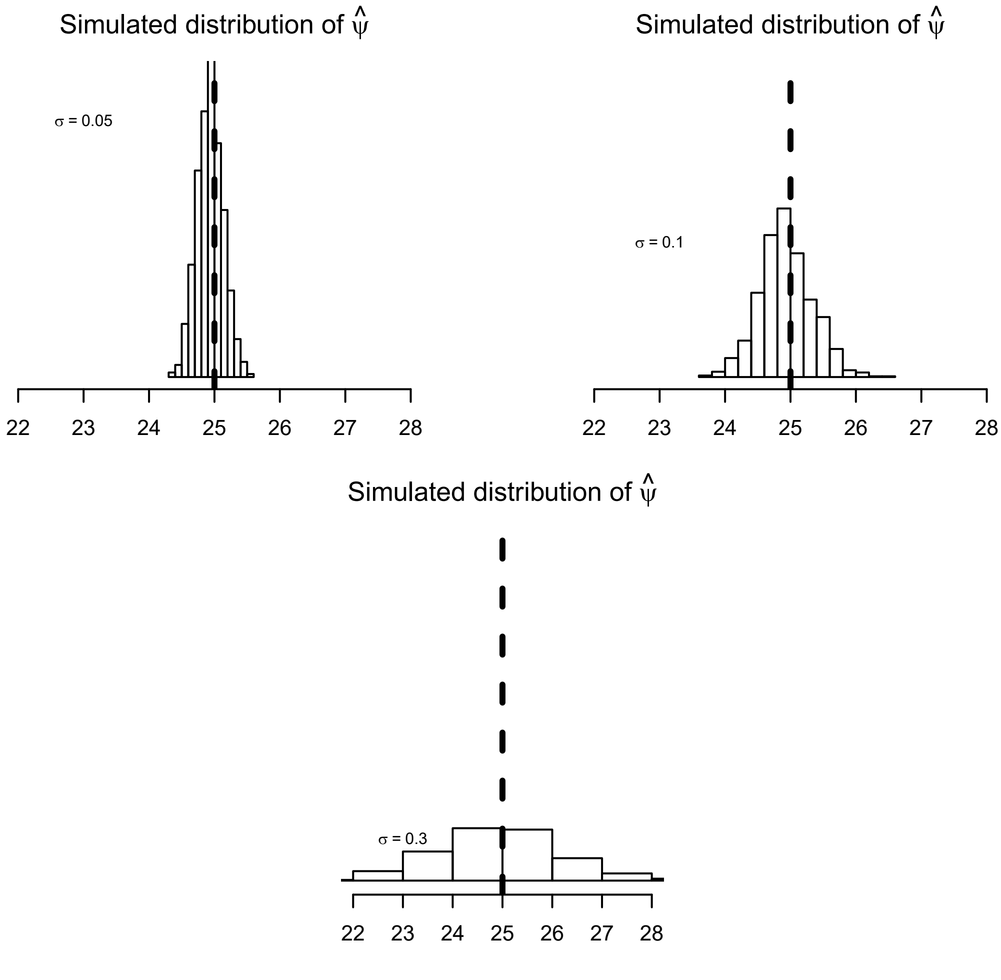

Figure 5.4: Simulated distribution of $\widehat{\psi}$ in E_{max} model with $n = 50$, $\psi = 25$, $\beta_{incr} = 2$, $\gamma = [n/2]$ and $\sigma = 0.05$, 0.1, 0.3.

is lowest and it is the highest for $\psi = [3n/4]$. For the other choices of n, we have the highest coverage for the lowest $\psi = [n/4]$. Similar behaviour was also visible

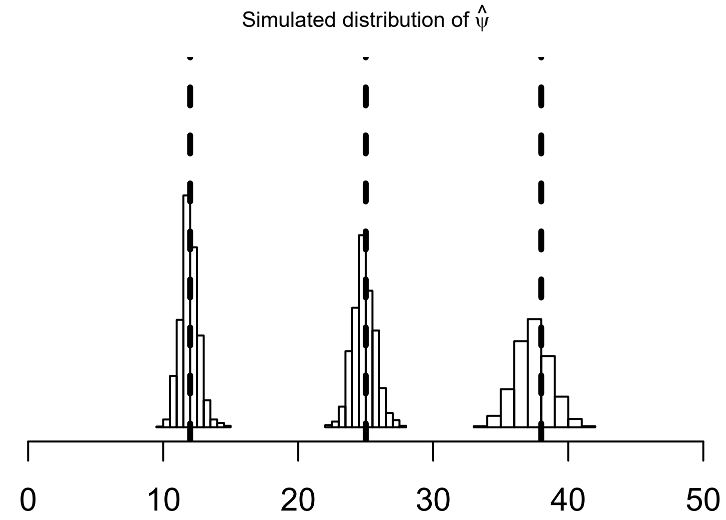

Figure 5.5: Simulated distribution of $\widehat{\psi}$ in E_{max} model for $n = 50, \psi = [n/4], [n/2], [3n/4]$, $\beta_{incr} = 5$, $\gamma = 25$ and $\sigma = 0.1$.

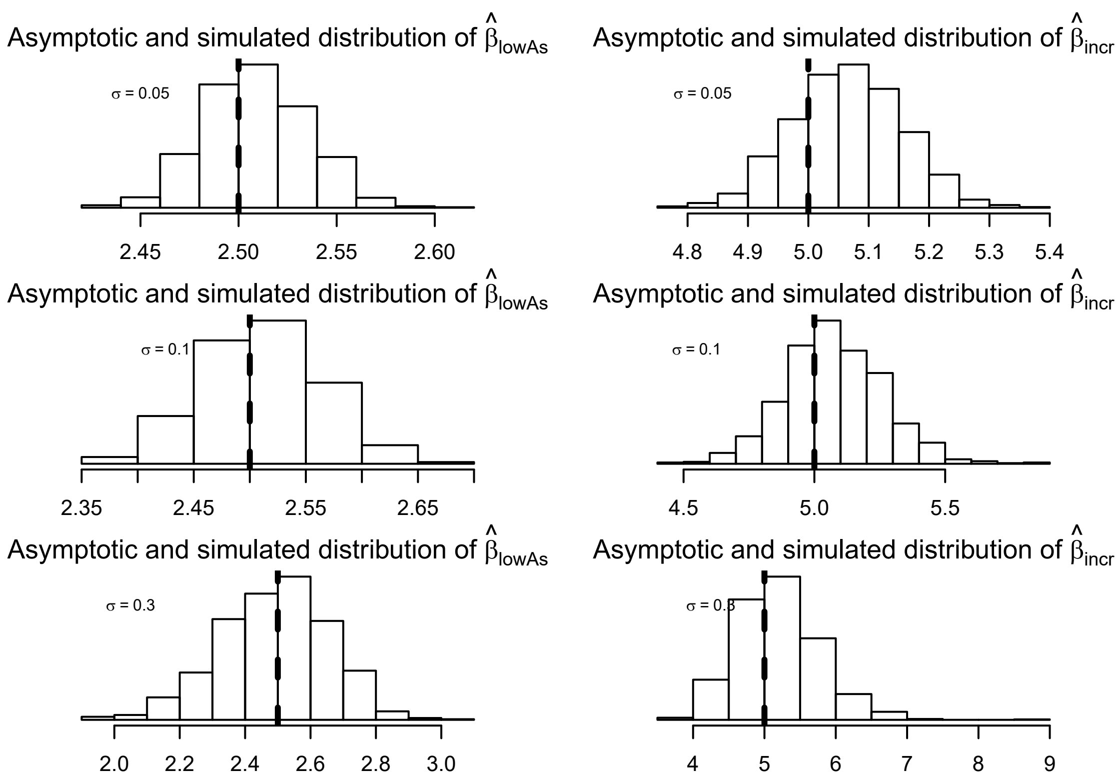

Figure 5.6: Simulated distribution of β_{lowAs} and β_{incr} in E_{max} model for with $n = 50$, $\psi = 25$, $\beta_{incr} = 5$, $\gamma = 25$ and $\sigma = 0.05, 0.1, 0.3$. The β_{lowAs} was calculated such that the mean value of the response after the change-points equals 5.

in Figure 5.5 where the estimators were less dispersed for the lowest ψ.

For the right-sided confidence interval, the results were similar as for the two-sided interval, see Table 5.4.

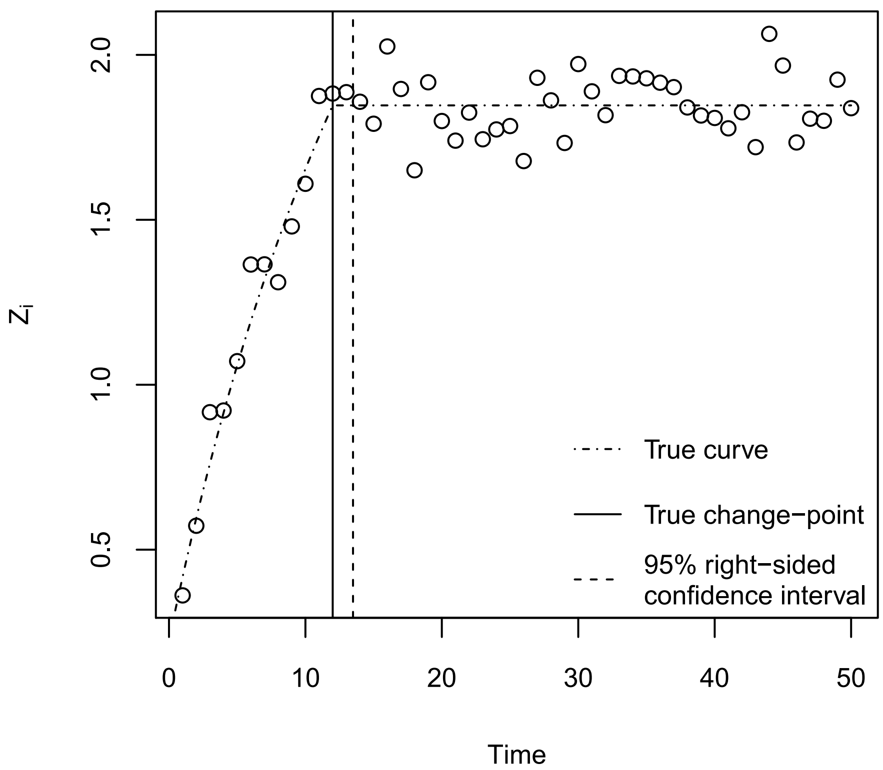

Figure 5.7: E_{max} model with 95% right-sided confidence interval for ψ for $n = 50$, $\psi = 25$, $\gamma = 25, \beta_{incr} = 5$, $\sigma = 0.1$, $\beta_{lowAs} = 0.2$ and $B = 1000$.

n	ψ	**Coverage bootstrap CI** [%]	**Mean length bootstrap CI**
	6	83.0	1.09
$n = 25$	12	88.9	1.38
	19	91.1	1.98
	12	94.7	1.56
$n = 50$	25	92.8	2.07
	38	91.8	2.95
	25	95.8	2.11
$n = 100$	50	93.6	2.96
	75	94.6	4.18

Table 5.3: Coverage and average length of a two-sided confidence interval (CI) for ψ in E_{max} model based on bootstrap for $\beta_{incr} = 3$, $\gamma = \psi$, $\sigma = 0.03$. The β_{lowAs} was calculated such that the mean value of the response after the change-points equals 3.

n	ψ	**Coverage bootstrap CI** [%]	**Mean distance** $c_U - \psi$ **bootstrap**
	6	79.9	0.41
$n = 25$	12	87.3	0.53
	19	90.1	0.79
	12	95.3	0.66
$n = 50$	25	93.4	0.84
	38	90.5	1.13
	25	96.1	0.87
$n = 100$	50	92.9	1.21
	75	92.2	1.70

Table 5.4: Coverage of a right-sided confidence interval (CI) for ψ and average distance $c_U - \psi$ in E_{max} model based on bootstrap for $\beta_{incr} = 3$, $\gamma = \psi$, $\sigma = 0.03$. The β_{lowAs} was calculated such that the mean value of the response after the change-points equals 3.

6. Model misspecification

With real data, it is often difficult to correctly specify the model that the data come from and whether the assumption of homoscedasticity of random errors is fulfilled. In this chapter, we discuss a situation when the model is incorrectly specified and we run simulations to find out how it influences the coverage of confidence intervals. We use results from previous chapters about *PoSt* models and we also address the wrong assumption of heteroscedasticity or homoscedasticity. We present two scenarios, one where we assume more complicated model than the data come from and second with less complex model than the true model.

For each scenario, we describe the assumed and the true model and assumptions concerning variances of the random errors. We run simulations and calculate the coverage of the confidence intervals for the change-point ψ and we discuss possible issues of misspecifying the model with real data.

In whole chapter, we assume we have m_i measurements at each $i = 1, \ldots, n$ from given model and we set $m_i = 20$. By assuming this, we can estimate the variance at each time i. We set $n = 50$, number of repetitions $N = 1000$, number of bootstrap samples $B = 1000$ and we try different locations of the change-point, $\psi = [n/4]$, $[n/2]$, $[3n/4]$.

We run simulations with $\sigma = 0.03$ and under heteroscedasticity, we set

$$\sigma_i = \sigma \left(1 + 1.5 \cdot \mathbb{1}_{[i \leq \psi]}\right), \quad i = 1, \ldots, n, \tag{6.1}$$

where $\mathbb{1}_{[x]}$ denotes the indicator function. This represents the situation, when the variance is bigger until the manufacturing line stabilises, i.e. before the change-point ψ.

During the estimation, we estimate the variance σ_i^2 at time i by the sample variance and we construct the weight matrix $\widehat{\mathbb{W}}$ as discussed in Chapter 4. This matrix is then used in estimation of the change-point and other parameters and in the bootstrap procedure.

Under the homoscedasticity assumption, we assume variances are the same for all $i = 1, \ldots, n$ and we set $\sigma = 0.03$. It can be estimated by the „pooled“ estimator of variance similarly as in Hlávka and Hušková [2017]. Denoting $\widehat{\sigma}_i^2$ the sample variance at time i, the pooled estimator is defined as

$$\widehat{\sigma}^2_{pooled} = \frac{\sum_{i=1}^{n} (m_i - 1)\, \widehat{\sigma}_i^2}{\sum_{i=1}^{n} (m_i - 1)}.$$

In each scenario, we generate data from the true model with the correct assumption of the variance structure (homoscedasticity or heteroscedasticity), we estimate unknown parameters using the (incorrect) assumed model and variance structure and we construct two-sided and right-sided confidence interval for ψ using bootstrap in the assumed model and variance structure. We also construct the confidence intervals using the true model for comparison.

In E_{max} model (2.16), the mean value of the output is increasing with $\beta_{incr} > 0$ and decreasing with $\beta_{incr} < 0$ up to the change-point ψ. Inversely, in a linear *PoSt* model, the mean value is increasing with $\beta_1 < 0$ and decreasing with $\beta_1 > 0$, similarly in quadratic *PoSt* model. To keep the same trend of the output variable, we manually change the sign of β_{incr} in E_{max} model when generating data and during estimation. When we set $\beta_1 = b$, and generate the data using E_{max} model, they will be generated using $\beta_{incr} = -b$. Similarly during estimation, estimating using E_{max} model on data with decreasing trend will resolve in $\widehat{\beta}_{incr} > 0$ by estimating the parameter β_{incr} and then changing the sign.

6.1 Overspecified model

In this section, we deal with a situation, when the assumed model for the data is more complicated than the true model from which the data are.

Let the true model for the data $Z_{i,j}, i = 1,, \ldots, n, j = 1, \ldots 20$ be a linear *PoSt* model with homoscedastic random errors, i.e.

$$Z_{ij} = \beta_0 + \beta_1 \left(\frac{\psi - i}{n}\right)^+ + e_{ij},$$

and $\operatorname{var} e_i = \sigma^2$ is same for all $i = 1, \ldots, n$. We set $\boldsymbol{\beta} = (2, 2)^\top$ and $\sigma = 0.03$.

In this setup, we try constructing the confidence interval assuming heteroscedasticity in linear, quadratic and E_{max} model. We also add the coverage for the true model and homoscedasticity.

One can expect the overspecification will not have such a big impact on the coverage as in the underspecification scenario. Results are summarised in Table 6.1.

With correctly specified model, i.e. linear *PoSt* model with homoscedastic errors, the coverage is around 95% and it is the highest for $\psi = [3n/4]$. When heteroscedasticity is assumed in the linear model, the coverage is a bit lower, around 92%, and slightly higher for the right-sided interval compared to the two-sided interval.

When assuming quadratic model, the coverage decreases to values around 88% for the two-sided interval and 92% for the right-sided interval.

The situation changes when assuming the E_{max} model. The coverage for $\psi = 12$ is only 40% and it is zero for the other locations of ψ. For the right-sided confidence interval, the coverage is 100%. This happens because by assuming E_{max} model and heteroscedasticity, we tend to have $\widehat{\psi} > \psi$, see the left side of Figure 6.1. This is because $\widehat{\psi}$ is found by maximizing a function (2.17) and the function has its maximum at larger values than ψ when using data from the homoscedastic linear model, see the right side of Figure 6.1 plotted for $\psi = 25$. Therefore the two-sided confidence interval does not cover true change-point ψ and the upper limit c_U of the right-sided confidence interval exceeds ψ. For $\psi = 25$, the average distance $c_U - \psi$ equals 2.3 and for $\psi = 38$ it equals 5.3 therefore the „bias“ of the estimator increases with increasing ψ. The situation is similar when assuming homoscedastic E_{max} model.

In general, assuming heteroscedasticity when homoscedasticity holds brings slightly lower coverage of the confidence intervals. When more complex polynomial model (the quadratic model) is assumed, the coverage drops a bit more to approximately 92% coverage of the right-sided interval. The problem occurs, when we incorrectly assume the data are from E_{max} model since the estimator $\widehat{\psi}$ tends to overestimate ψ because of the shape of maximised function.

Assumed model	Assumed variace structure	Change-point	Both-sided CI coverage [%]	Right-sided CI coverage [%]
		12	95.2	94.0
linear	homoscedastic	25	95.5	94.5
		38	96.2	95.5
		12	91.8	92.3
linear	heteroscedastic	25	91.7	92.1
		38	91.8	93.5
		12	88.7	90.0
quadratic	heteroscedastic	25	85.4	92.1
		38	91.1	92.0
		12	40.4	100.0
E_{max}	heteroscedastic	25	0.0	100.0
		38	0.0	100.0

Table 6.1: Coverage of bootstrap 95% confidence intervals for ψ in the correct and in the overspecified model, the true model is homoscedastic linear *PoSt* model with $n = 50$, $\sigma = 0.03$, $\beta_0 = \beta_1 = 2$.

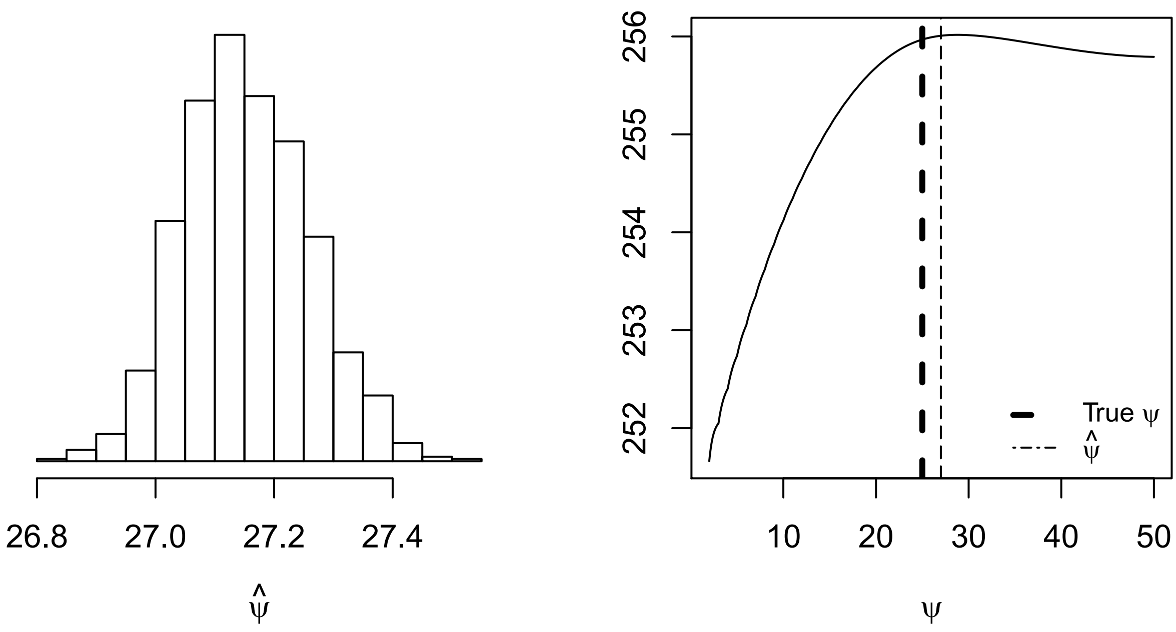

Figure 6.1: Estimation when assuming heteroscedastic E_{max} model and the true model is homoscedastic linear with $\psi = 25$, $n = 50$, $\beta_0 = \beta_1 = 2$. Left: Histogram of simulated $\widehat{\psi}$. Right: Function to be maximised when estimating ψ for one simulated data.

6.2 Underspecified model

In this section, we deal with a situation, when the data are assumed to be from a simpler model than they really are.

Let true model for the data $Z_{i,j}, i = 1, , \dots, n$, $j = 1, \dots 20$ be a quadratic *PoSt* model

$$Z_{ij} = \beta_0 + \beta_1 \left(\frac{\psi - i}{n}\right)^+ + \beta_2 \left(\left(\frac{\psi - i}{n}\right)^+\right)^2 + e_{ij}$$

with heteroscedastic random errors with $\operatorname{var} e_{ij} = \sigma_i^2$ defined by (6.1) and parameters $\beta_0 = \beta_1 = \beta_2 = 2$.

With this data, we estimate parameters and we construct bootstrap confidence intervals assuming homoscedasticity in linear, quadratic and E_{max} model. We can expect the coverage will be lower than in overspecified scenario since we assume simpler model than the true model.

The results are summarised in Table 6.2. When the true model is assumed, the coverage is also lower than 95%. It is possible, that in some cases the change-point was incorrectly estimated because of more dispersed data before the change-point and the confidence interval, constructed from samples generated using this „shifted" estimators, did not cover the true change-point. Smaller values σ_i^2 or larger β_1, β_2 would probably make the coverage higher. The coverage of the right-sided interval is higher, around 90%.

With incorrectly specified model, the coverage varies and estimators are in general not reliable.

When quadratic model with homoscedasticity is assumed, the coverage is lower, we have 74% coverage for $\psi = 12$, 50.3% for $\psi = 25$ and 93.9% for $\psi = 38$ for the two-sided confidence interval. For the right-sided interval, the coverage is higher for $\psi = 12$, 25 compared to two-sided interval.

When assuming linear model, the situation changes. For $\psi = 12$ the coverage is $< 25\%$ and it is 0 for the other chosen ψ for both types of intervals. This happens, because by incorrectly assuming linear model we tend to have $\widehat{\psi} < \psi$, see the left side of Figure 6.2, since the maximum of maximised function (2.2) for linear *PoSt* model is not the true change-point ψ, see the right side of Figure 6.2. Therefore, both the two-sided and the right-sided confidence intervals do not cover the true ψ.

The situation is again different for the E_{max} model. We get coverage higher than 75% for $\psi = 12$, 25 and the coverage drops to only 0.1% for $\psi = 38$. Again, the coverage of the right-sided interval is much higher, we often obtain $\widehat{\psi} > \psi$ similarly as in the overspecification scenario.

When we are not sure about the assumption of homoscedasticity, it is therefore safer to assume heteroscedasticity since otherwise the coverage is not reliable. Also, we should not assume simpler model since it yields in very low coverage.

Assumed model	Assumed variace structure	Change-point	Both-sided CI coverage [%]	Right-sided CI coverage [%]
quadratic	heteroscedastic	12	87.4	90.4
		25	84.3	90.5
		38	91.2	90.4
linear	homoscedastic	12	20.6	16.4
		25	0.0	0.0
		38	0.0	0.0
quadratic	homoscedastic	12	74.0	79.4
		25	50.3	89.2
		38	93.9	92.4
emax	homoscedastic	12	87.5	91.8
		25	77.9	99.7
		38	0.1	100.0

Table 6.2: Coverage of bootstrap 95% confidence intervals for ψ in the correct and in the underspecified model, the true model is heteroscedastic quadratic *PoSt* model with $n = 50$, $\beta_0 = \beta_1 = \beta_2 = 2$ and σ_i given by (6.1) with $\sigma = 0.03$.

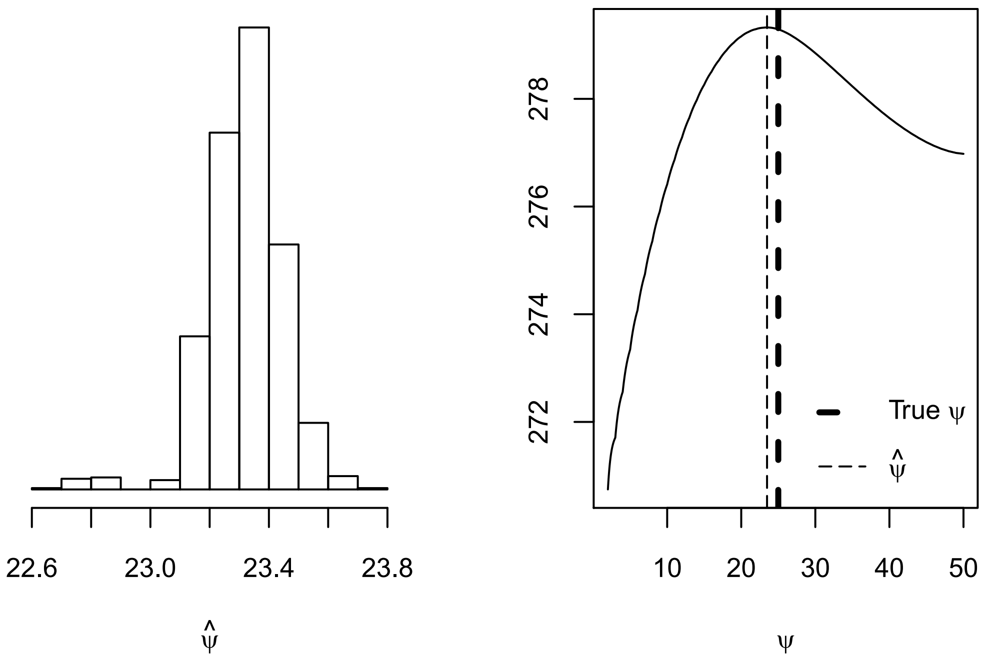

Figure 6.2: Estimation when assuming homoscedastic linear *PoSt* model and the true model is heteroscedastic quadratic with $\psi = 25$, $n = 50$, unknown β_0 and $\beta_0 = \beta_1 = \beta_2 = 2$. Left: Histogram of simulated $\widehat{\psi}$. Right: Function to be maximised when estimating ψ for one simulated data.

Conclusion

In the thesis, we dealt with gradual change model, mostly in point-of-stabilisation (*PoSt*) context. First, we reviewed available methods, we briefly discussed testing the presence of the change-point and in more detail, we discussed the estimation in gradual change models with polynomial trend and also in more general setups, such as panel data, in Chapter 1.

In Chapter 2, we first introduced least squares estimators of unknown parameters in gradual change model with polynomial trend as in Hušková [1998] and we stated their asymptotic results. In Section 2.2 we introduced the model with quadratic trend and „reversed“ time ordering which is used in *PoSt* context. For the quadratic model, we stated estimators of unknown parameters and asymptotic results derived in Jarušková [2001] where we extended used theorems by assuming general variance σ. In Section 2.3, we introduced the E_{max} model in its general form used e.g. in dose-response studies in MacDougall [2006] and we modified the model by including the change-point and we derived estimators similarly as in linear change model.

In Chapter 3, we introduced the linear *PoSt* model, motivated by estimating the point-of-stabilisation in continuous manufacturing process. We made use of estimators derived in previous section about linear and general polynomial trend. When deriving the asymptotic distribution, we translated results from Hušková [1998] into reversed time settings used in *PoSt* context. Next, we used derived results to construct confidence intervals for the change-point, we discussed the connection to testing and we explained how they can be used in practice to verify the stability of the production process. In Section 3.5 we visualised the difference between the asymptotic distribution and the simulated distribution for various sample sizes n, locations of the change-point ψ and variance σ. Here, we found out properties of the change-point estimator $\widehat{\psi}$ depend on the variance σ and also on the location of the true change-point, the best approximation is for the highest value $\psi = [3n/4]$ among all selected locations of ψ. We calculated the coverage of two-sided and right-sided confidence interval for ψ based on asymptotic results and on bootstrap. For higher n, the coverage was similar for all selected locations of ψ and for $n = 100$, the coverage was around 94.5%. for the asymptotic confidence interval. Using bootstrap, the coverage was in most cases slightly higher and the bootstrap confidence interval was on average also a bit wider. For the right-sided interval, the asymptotic confidence interval had slightly higher coverage than the bootstrap interval.

Therefore, when assuming linear *PoSt* model and calculating right-sided confidence interval, we would recommend to use the interval based on the asymptotic distribution. On the other hand, for the change-point located early in the data (e.g. $\psi = [n/4]$), bootstrap provides slightly better coverage as we saw in Table 3.2 for $n = 25, 100$.

In Chapter 4, we showed how to modify the estimation process when homoscedasticity is not fulfilled by using multiple measurements and estimating the variance σ_i^2 for each time i. We described the bootstrap procedure to get the confidence interval for the change-point since in this situation we do not have any asymptotic results.

We then aimed at nonlinear *PoSt* models in Chapter 5, namely quadratic and E_{max} model. We discussed the construction of confidence intervals, which can be used in practice, and we visualised the simulated and the asymptotic distribution. For the quadratic model, the approximation was similar as in linear *PoSt* model.

We calculated the coverage of confidence intervals as in the linear *PoSt* model. For the quadratic model, the difference in coverage for different locations of ψ was more evident. For $\psi = [n/4]$ the coverage was the lowest for all chosen n and the coverage increased with increasing ψ. Here, bootstrap confidence intervals had slightly lower coverage similarly as in linear *PoSt* model. For right-sided bootstrap interval coverage, results were similar as for two-sided confidence interval.

Therefore, when assuming quadratic model, we would again recommend using the confidence interval based on asymptotic results. On the other hand, bootstrap confidence interval provides better coverage in the case of „early“ change-point, in our case $\psi = [n/4]$, similarly as in linear *PoSt* model.

Next, we visualised the simulated distribution for the E_{max} model in *PoSt* context together with true values of parameters. We found out that estimators of the change-point are less dispersed for lower values of θ, compared to linear and quadratic *PoSt* model. The coverage of bootstrap confidence intervals for ψ in E_{max} model was slightly below 95% and it was the highest for $\psi = [n/4]$.

In the end in Chapter 6, we discussed what happens when the true model is not correctly specified and when incorrect assumption of homoscedasticity or heteroscedasticity is done. When assuming more complex model than the true model and hetroscedasticity when homoscedasticity is fulfilled, the coverage for ψ was slightly lower, but still mostly above 91% for the right-sided confidence interval, see Table 6.1. When E_{max} model was assumed, the coverage of the two-sided interval was only 40% for $\psi = [n/4]$ and none of bootstrap confidence intervals covered the true ψ for $\psi = [n/2]$, $[3n/4]$. For the right-sided interval, we had 100% coverage since the estimated change-points tend to overestimate the true ψ because of the shape of maximised function, see Figure 6.1.

When assuming less complex model and homoscedasticity while heteroscedastic quadratic model holds, the results were not reliable, see Table 6.2. When assuming correct model and homoscedasticity, the coverage of the right-sided confidence interval was around 85%. When assuming linear homoscedastic model, the coverage was below 20% because estimators $\widehat{\psi}$ tend to underestimate the true ψ, see Figure 6.2. When assuming the E_{max} model, the situation was similar as in overspecification scenario, we again tend to have $\widehat{\psi} > \psi$ and the coverage of the right-sided confidence interval was mostly higher than 95%.

Therefore, when one is not sure about homoscedasticity, we would recommend to assume heteroscedasticity, which brings only small drop in coverage if homoscedasticity holds. Also, assuming a simpler model yields in very low coverage, we would therefore recommend assuming model with more complex polynomial trend. On the other hand, misspecifying the E_{max} model results in „biased“ change-point estimator because of the shape of maximised function.

Bibliography

J. Anděl. *Základy matematické statistiky.* Druhé vydání. Matfyzpress, Praha, 2007.

R. P. Brent. *Algorithms for minimization without derivatives.* Courier Corporation, 2013.

M. Döring. Asymmetric cusp estimation in regression models. *Statistics*, 49(6): 1279–1297, 2015.

R. Fletcher and C. M. Reeves. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

Z. Hlávka and M. Hušková. Two-sample gradual change analysis. *REVSTAT–Statistical Journal*, 15(3):355–372, 2017.

M. Hušková. Estimators in the location model with gradual changes. *Comment. Math. Univ. Carolin.*, 39(1):147–157, 1998.

M. Hušková. Gradual changes versus abrupt changes. *Journal of Statistical Planning and Inference*, 76(1-2):109–125, 1999.

M. Hušková and J. Steinebach. Limit theorems for a class of tests of gradual changes. *Journal of statistical planning and inference*, 89(1-2):57–77, 2000.

D. Jarušková. Change-point estimator in gradually changing sequences. *Comment. Math. Univ. Carolinae*, 39:551–561, 1998a.

D. Jarušková. Testing appearance of linear trend. *Journal of Statistical Planning and Inference*, 70(2):263–276, 1998b.

D. Jarušková. Change-point estimator in continuous quadratic regression. *Comment. Math. Univ. Carolin*, 42(4):741–752, 2001.

S. L Lee, T. F O'Connor, X. Yang, C. N. Cruz, S. Chatterjee, R. D. Madurawe, Ch. M. V. Moore, X. Y. Lawrence, and J. Woodcock. Modernizing pharmaceutical manufacturing: from batch to continuous production. *Journal of Pharmaceutical Innovation*, 10(3):191–199, 2015.

J. MacDougall. *Analysis of Dose–Response Studies—Emax Model*, pages 127–145. Springer, New York, 2006.

Š. Rusá. Panelová data a detekce změn. *Diplomová práce, MFF UK*, 2015.

S. D. Schaber, D. I. Gerogiorgis, R. Ramachandran, J. M.B. Evans, P. I. Barton, and B. L. Trout. Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. *Industrial & Engineering Chemistry Research*, 50(17):10083–10092, 2011.

A.W. Van der Vaart. *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. doi: 10.1017/CBO9780511802256.