Statistical Inference 2 Summer Term 2025

Lecture: Prof. Dr. Göran Kauermann
Exercises: Sergio Buttazzo, Jan Anders
Tutorial: Eugen Gorich

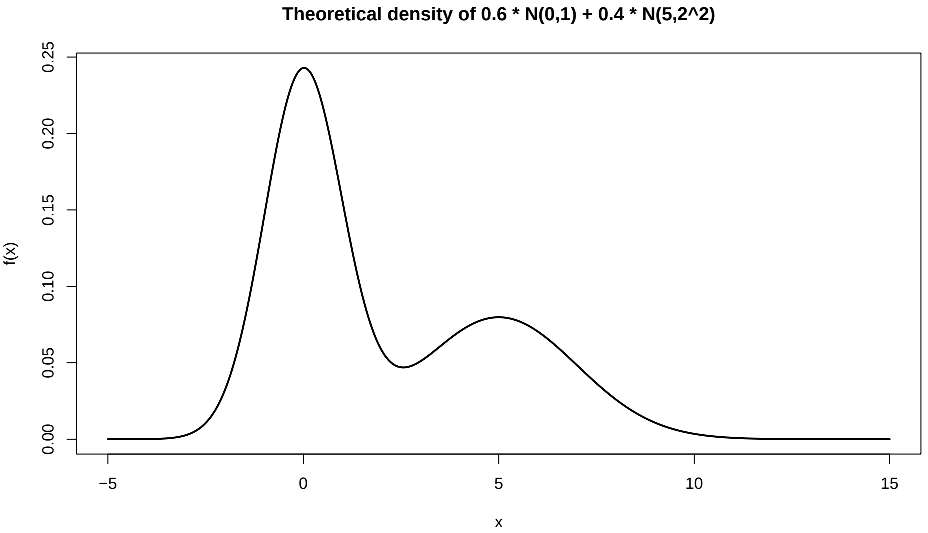
Exercise Sheet 2

Exercise 2.1 - Bootstrapping

Let $X_1, \dots, X_{120}$ be an i.i.d. sample from the two-component normal mixture

$$X \sim 0.6\, N(0,1) + 0.4\, N(5, 2^2).$$

That means the distribution of X looks like this:



The theoretical interquartile range of this mixture is $Q(0.75) - Q(0.25) = 4.581$. In this exercise, we want to use nonparametric bootstrapping to study the sampling distribution of the sample interquartile range

$$I = Q_n(0.75) - Q_n(0.25),$$

where $Q_n(p)$ is the empirical p-quantile.

a) Generate one sample of size $n = 120$ (set seed 123) from the mixture distribution (i.e. generate each sampled point from the first component with probability 0.6 and from the second with probability 0.4) and compute $I_{\text{obs}} = Q_n(0.75) - Q_n(0.25)$.

b) From the empirical distribution generated in (a), draw $B = 500$ bootstrap samples (set seed 1234) and compute I_b^* for each. Compute the bootstrap standard error $\widehat{\text{SE}}_{\text{boot}}$ and the skewness of $\{I_b^*\}$.

c) Plot a histogram of $\{I_b^*\}$ with an overlaid kernel density estimate.

d) Construct a 95% percentile bootstrap CI and a normal-approximation CI for the true IQR. Compare their width.

e) Discuss: How does the mixture's skew affect the bootstrap distribution and the reliability of the normal CI? What would change if n were larger?

Exercise 2.2 - Bayesian Reasoning

An industrial plant is evaluating four different predictive maintenance models (M_1 through M_4). Each model outputs a daily probability of machine failure. Over the past 30 days, each model provided a probability estimate for whether a key machine would fail, and the actual outcomes (failure or not) are recorded as $x_1, \ldots, x_{30}$, where $x_t = 1$ indicates a failure and $x_t = 0$ indicates normal operation. We will assume in this exercise that the probability of a failure does not depend on t. The historical reliability (from more than 30 days ago) of the models is encoded in the following prior probabilities:

$$P(M_1) = 0.3, \quad P(M_2) = 0.35, \quad P(M_3) = 0.2, \quad P(M_4) = 0.15$$

In `Ex_3_2.RData`, available on Moodle, you are given the predicted probabilities π_{it} of failure on each day $t = 1, \ldots, 30$ for each model M_i, as well as the observed failure outcomes $x_1, \ldots, x_{30}$ and the probabilities $\pi_{i,\text{new}}$ that will be relevant in (b).

(a) Compute the posterior probability of each model given the observed outcomes, using Bayes' theorem. Express the likelihood for each model as the product of Bernoulli likelihoods over the 30 days:

$$P(x_1, \ldots, x_{30} | M_i) = \prod_{t=1}^{30} \pi_{it}^{x_t} (1 - \pi_{it})^{1-x_t}$$

Which model is most plausible after observing the data?

(b) Each model provides a probability forecast $\pi_{i,\text{new}}$ for failure tomorrow (See `$pi_new` of the RData object provided). Compute the posterior predictive probability of failure tomorrow $\hat{\pi} = P(\text{failure tomorrow}|\text{data})$, as the sum of $\pi_{i,new}$, weighted by the posterior reliability of the models from (a).

(c) Suppose instead that in the 30-day window, failures had occurred on 20 days. In general, how would this affect the posterior model probabilities and the posterior predictive probability of failure tomorrow? Comment on which models become more plausible and how this affects $\hat{\pi}$. For this, you may first need to look at the distribution of each model's predictions.

(d) The plant must decide every day whether to perform preventive maintenance. The costs for this can be described as a loss function $L(\theta, d)$, where θ is a binary variable (F = failure, NF=no failure) indicating if failure actually (would have) occurred and d is the decision that was made (M = maintenance, NM = no maintenance).

$$L(\theta, d) = \begin{cases} 0 & \text{if failure would have occurred but maintenance was done (F, M)} \\ q & \text{if failure occurs and no maintenance (cost of production stop) (F, NM)} \\ 1 & \text{if no failure and maintenance done (NF, M)} \\ 0 & \text{if no failure and no maintenance (NF, NM)} \end{cases}$$

Assume $q = 20$ and $\mathbb{E}(\hat{\pi}) = \mathbb{E}(\theta)$. Compute the Bayes risk, i.e. $\mathbb{E}(L(\theta, d) \mid d)$ of each decision d and determine the optimal action given $\hat{\pi} = P(\text{failure tomorrow}|\text{data})$ for any given day.

(e) Determine the critical value q^* (depending on π) at which the plant is indifferent between maintaining and not maintaining. What does this imply about risk sensitivity?

Exercise 2.3 - Conjugate Priors

Let $X = (X_1, \dots, X_n)$ be an i.i.d. exponentially distributed random vector with parameter λ, that is $X_i \sim Exp(\lambda),\ i = 1, \dots, n$. We are interested in the distribution of the unknown parameter λ, and assume a gamma distribution for λ, so that $\lambda \sim Gamma(\alpha, \beta)$ and $X_i|\lambda \sim Exp(\lambda)$.

The density of the exponential distribution is given by: $f(x; \lambda) = \lambda e^{-\lambda x}$. For a gamma distributed variable Y, it holds that

$$f(y; \alpha, \beta) \propto y^{\alpha-1} e^{-\beta y}$$

with the expected value given by $\mathbb{E}(Y) = \frac{\alpha}{\beta}$.

(a) Explain the concept of a conjugate prior.

(b) Show that the gamma distribution is a conjugate prior for the exponential likelihood by computing the posterior distribution. State the setup and explain your steps.

(c) Your prior knowledge on λ can be captured by Gamma(2, 3). You have now collected a sample $x = (1, 1, 2, 1, 3, 1, 4, 6, 1, 1)$. Specify the updated posterior parameters α^* and β^* and compute the posterior mean $\mathbb{E}(\lambda^*)$.

(d) What are the advantages and disadvantages of using conjugate priors? Discuss how these properties affect the choice of prior distribution and the resulting posterior inference.