MATH562 – Fall 2025

Problem Set: Week 5

*1. A location-scale model is of the form $y_0 = \eta + \tau\, v_0$, where v_0 has a known distribution, with location and scale parameters $\eta \in \mathbb{R}$ and $\tau > 0$. A location estimator $m(y)$ and scale estimator $s(y)$, based on a random sample $y = (y_1, \ldots, y_n)$ with the y_i iid. distributed as y_0, are said to be *equivariant* if they satisfy

$$\begin{aligned} m(a + b\,y) &= m(a + b\,y_1, \ldots, a + b\,y_n) = a + b\,m(y_1, \ldots, y_n) = a + b\,m(y), \\ s(a + b\,y) &= s(a + b\,y_1, \ldots, a + b\,y_n) = b\,s(y_1, \ldots, y_n) = b\,s(y), \quad a \in \mathbb{R}, b > 0. \end{aligned}$$

(a) Show that $q_1(y, \eta) = \{m(y) - \eta\}/s(y)$ and $q_2(y, \tau) = s(y)/\tau$ are pivots, and explain how to use them to construct confidence intervals for η and τ.

(b) Deduce that pivots can be formed by taking

(i) $m_1(y) = \bar{y}$ and $s_1(y) = \left\{\sum_j (y_j - \bar{y})^2\right\}^{1/2}$,

(ii) $m_2(y) = \text{median}(y)$ and $s_2(y) = \text{IQR}(y)$, the interquartile range,

whatever the distribution of v_0. Briefly discuss the corresponding confidence intervals.

(c) Provide a pivot that could be used to construct a prediction interval for y_0 based on $y_1, \ldots, y_n$.

2. Find an α-level confidence interval for Y_+, when $Y_1, \ldots, Y_n, Y_+ \overset{i.i.d.}{\sim} \exp(\lambda)$. How does the interval change as $n \to \infty$. *Hint: if $E \sim \exp(1)$ then $2E \sim \chi^2_2$.*

3. (a) An experiment consists of observing the number of success y_1 in a fixed number n_1 of independent Bernoulli trials with unknown success probability $\theta \in (0, 1)$. Show that the corresponding density is

$$f(y_1 \mid \theta) = \binom{n_1}{y_1} \theta^{y_1} (1 - \theta)^{n_1 - y_1}, \quad y_1 \in \mathcal{S}_\infty = \{0, 1, \ldots, n_1\}.$$

If prior information on θ can be summarised by the *beta density*

$$f(\theta) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1 - \theta)^{b-1}, \quad 0 < \theta < 1, \quad a, b > 0,$$

show that the posterior density for θ given y_1 is

$$f(\theta \mid y_1) = \frac{\Gamma(n_1 + a + b)}{\Gamma(y_1 + a)\Gamma(n_1 - y_1 + b)} \theta^{y_1 + a - 1} (1 - \theta)^{n_1 - y_1 + b - 1}, \quad 0 < \theta < 1.$$

(b) Another experiment conducts independent Bernoulli trials until there are y_2 successes, at which point there have been n_2 trials. Show that the corresponding density is

$$f(n_2 \mid \theta) = \binom{n_2 - 1}{y_2 - 1} \theta^{y_2} (1 - \theta)^{n_2 - y_2}, \quad n_2 \in \mathcal{S}_\in = \{y_2, y_2 + 1, \ldots\}.$$

Without doing any calculations, write down the posterior density for θ based on the prior in (a).

(c) Show that if $y_1 = y_2$ and $n_1 = n_2$, then Bayesian inferences based on either of the two experiments will be identical, i.e., they do not take into account the different reference sets $\mathcal{S}_\infty$ and $\mathcal{S}_\in$.

(d) Consider testing the hypothesis that $\theta = \frac{1}{2}$ against the alternative that $\theta < \frac{1}{2}$. Explain why the respective significance levels for the experiments in (a) and (b) would be

$$\sum_{y=0}^{y_1} \binom{n_1}{y} 2^{-n_1}, \qquad \sum_{n=n_2}^{\infty} \binom{n-1}{y_2-1} 2^{-n},$$

and evaluate these when $n_1 = n_2 = 12$, $y_1 = y_2 = 3$. How does this compare with (c)?

Hint: Recall that for $\alpha > 0$ the gamma function is defined as $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u}\, du$, and that $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(n+1) = n!$ for $n \in \{1, 2, \ldots\}$, and $\Gamma(1/2) = \sqrt{\pi}$.

4. Consider the shoe data example from the Slides (Part 1).

   (a) Show that the average $\bar{D}$ has mean θ and variance $\sigma^2 = m^{-2} \sum_{j=1}^{m} c_j^2$.

   (b) Show that S^2 has mean σ^2 and hence can be used to estimate σ^2.

   (c) Extend this discussion to a balanced design in which m is even, $I_j = \pm 1$ and $\sum_{j=1}^{m} I_j = 0$ but the allocation is otherwise completely at random. This ensures that materials A and B appear equally often on the left and right shoes.