MATH562 – Fall 2025

Problem Set: Week 13

*1. Let $T \sim \mathbb{N}(\mu, \sigma^2)$ with σ^2 known. Inference for the unknown μ proceeds in two stages. First, a two-sided test of the hypothesis that $\mu = 0$ is conducted. If this test is significant at level 2α, then a confidence interval $\mathcal{I}_{1-2\alpha}$ with limits $T \pm \sigma z_{1-\alpha}$ is computed. Show that if $\mu \geq 0$ then the coverage of this interval is

$$\frac{1 - \alpha - \Phi\{\max(z_\alpha, z_{1-\alpha} - \mu/\sigma)\}}{\Phi(z_\alpha + \mu/\sigma) + \Phi(z_\alpha - \mu/\sigma)},$$

and discuss how this depends on $\theta = \mu/\sigma$.

Solution:

The confidence interval is computed only if the test is significant, and this occurs if $|T| > \sigma z_{1-\alpha}$, an event with probability

$$\begin{aligned}
\Pr(T < -\sigma z_{1-\alpha}) + \Pr(T > \sigma z_{1-\alpha}) &= \Pr(Z < -z_{1-\alpha} - \mu/\sigma) + \Pr(Z > z_{1-\alpha} - \mu/\sigma) \\
&= \Pr(Z < z_\alpha - \mu/\sigma) + \Pr(Z < z_\alpha + \mu/\sigma) \\
&= \Phi(z_\alpha + \mu/\sigma) + \Phi(z_\alpha - \mu/\sigma),
\end{aligned}$$

where $Z = (T - \mu)/\sigma \sim \mathbb{N}(0, 1)$ and Φ and z_p are respectively the standard normal CDF and its p quantile, for $0 < p < 1$.

The confidence interval is computed only if $|T| > \sigma z_{1-\alpha}$, so the true coverage is the conditional probability

$$\Pr(\mu \in \mathcal{I}_{1-2\alpha} \mid |T| > \sigma z_{1-\alpha}) = \frac{\Pr(\{T - \sigma z_{1-\alpha} \leq \mu \leq T + \sigma z_{1-\alpha}\} \cap \{|T| > \sigma z_{1-\alpha}\})}{\Pr(|T| > \sigma z_{1-\alpha})}.$$

The numerator event here is

$$\{z_\alpha \leq Z \leq z_{1-\alpha}\} \cap (\{Z < z_\alpha - \mu/\sigma\} \cup \{Z > z_{1-\alpha} - \mu/\sigma\}) = \{z_\alpha \leq Z \leq z_{1-\alpha}\} \cap \{Z > z_{1-\alpha} - \mu/\sigma\},$$

because $\mu > 0$. This is $\{z_\alpha \leq Z \leq z_{1-\alpha}\}$ if $z_{1-\alpha} - \mu/\sigma < z_\alpha$, but otherwise is $\{z_{1-\alpha} - \mu/\sigma \leq Z \leq z_{1-\alpha}\}$, and hence has probability

$$\Pr\{\max(z_\alpha, z_{1-\alpha} - \mu/\sigma) \leq Z \leq z_{1-\alpha}\} = \Phi(z_{1-\alpha}) - \Phi\{\max(z_\alpha, z_{1-\alpha} - \mu/\sigma)\},$$

as required.

When $\mu = 0$ the coverage is zero, since the interval is computed only when the hypothesis $\mu = 0$ is rejected, which is equivalent to the interval not containing μ. When μ is small and positive, the interval is again unlikely to contain μ, because the event $|T| > \sigma z_{1\alpha}$ pushes T outside the upper rejection limit $\sigma z_{1\alpha}$. As μ increases the interval is more likely to contain μ, because the event $|T| > \sigma z_{1\alpha}$ corresponds increasingly to $T > \sigma z_{1\alpha}$. Finally there is a cusp in the probability when $z_\alpha = z_{1-\alpha} - \mu/\sigma$, i.e., $\mu/\sigma = 2z_{1-\alpha}$, after which only the denominator probability increases, thereby reducing the coverage to its correct value of $1 - 2\alpha$.

2. Randomisation for selective inference on θ based on $T \sim \mathbb{N}(\theta, 1)$ and an independent $W \sim \mathbb{N}(0, 1)$ results in $U = T + pW$ and $V = T - W/p$.

   (a) Find the distribution of U, and that of T conditional on $U = u$. **Solution**: Clearly U is normal with mean θ and variance $1 + p^2$, and $\text{Cov}(T, U) = \text{Var}(T) = 1$, so T is conditionally normal with mean and variance

   $$\text{E}(T \mid U = u) = \theta + (u - \theta)/(1 + p^2), \quad \text{Var}(T \mid U = u) = 1 - 1^2/(1 + p^2) = p^2/(1 + p^2).$$

(b) Show that the expected information for θ can be decomposed as

$$\imath(\theta) = \mathrm{E}_U\left\{-\frac{\mathrm{d}^2 \log f(U;\theta)}{\mathrm{d}\theta^2}\right\} + \mathrm{E}_U\left[\mathrm{E}_{T|U}\left\{-\frac{\mathrm{d}^2 \log f(T \mid U;\theta)}{\mathrm{d}\theta^2}\right\}\right],$$

and find these terms in the situation above. Hence deduce that randomisation can be viewed as partitioning the overall information between the two phases of inference. **Solution**: We have $f(t,u;\theta) = f(u;\theta)f(t \mid u;\theta)$, so taking logs, differentiation and then taking expectations will lead to the given expression for $\imath(\theta)$.

In the particular case of the normal model, and ignoring additive constants, the two terms of the log likelihood are

$$-\frac{(u-\theta)^2}{2(1+p^2)}, \quad -\frac{\{t-\theta-(u-\theta)/(1+p^2)\}^2}{2p^2/(1+p^2)},$$

and differentiation of these expressions twice with respect to θ gives

$$-\frac{1}{1+p^2}, \quad -\frac{p^2}{1+p^2},$$

so the two terms in the information decomposition sum to the overall information $\imath(\theta) = 1$. If p is small, then the first term, corresponding to U, comprises almost all the overall information, but that for inference (the second term) is small, and conversely when $p \approx 1$.

3. To investigate how randomisation for dealing with selection might be applied to Poisson variables, suppose that Y has a Poisson distribution with mean θ, and let $X = (X_1, \ldots, X_K)$ be a multinomial variable with denominator y and probability vector $(p_1, \ldots, p_K)$.

(a) Given that the joint moment-generating function (MGF) of X given that $Y = y$ is

$$\mathrm{E}\left\{\exp\left(\sum_{k=1}^K t_k X_k\right) \middle| Y = y\right\} = \left(\sum_{k=1}^K p_k e^{t_k}\right)^y, \quad t_1, \ldots, t_K \in \mathbb{R},$$

show that after marginalisation over Y the components of X are independent Poisson variables with means $p_1\theta, \ldots, p_K\theta$. (Recall that a Poisson variable with mean λ has MGF $\exp\{\lambda(e^t - 1)\}$.)

Solution: The marginal MGF of X is

$$\begin{aligned}
\mathrm{E}_Y\left[\mathrm{E}\left\{\exp\left(\sum_{k=1}^K t_k X_k\right) \middle| Y = y\right\}\right] &= \sum_{y=0}^\infty \left(\sum_{k=1}^K p_k e^{t_k}\right)^y \theta^y e^{-\theta}/y! \\
&= \exp\left\{-\theta + \theta \sum_{k=1}^K p_k e^{t_k}\right\} \\
&= \prod_{k=1}^K \exp\left\{p_k\theta(e^{t_k} - 1)\right\},
\end{aligned}$$

which is the MGF of K independent Poisson variables with with means $p_1\theta, \ldots, p_K\theta$. Here we wrote $\theta = \theta(P_1 + \cdots + p_K)$.

(b) Discuss the use of randomisation for selective inference when the data consist of independent Poisson variables $Y_1, \ldots, Y_n$ with means $\theta_1, \ldots, \theta_n$. **Solution**: We aim to generate a set of Poisson variables $Y_1^*, \ldots, Y_n^*$ to be used for selection and an independent set of Poisson variables $Y_1^\dagger, \ldots, Y_n^\dagger$ to be used for inference.

The result from (a) suggests that we might choose $p \in (0,1)$ and then generate $Y_j^* \sim B(Y_j, p)$, which will be independent with means $p\theta_j$, also taking $Y_j^\dagger = Y_j - Y_j^*$, which will be independent (and independent of the Y_j, unconditionally), with means $(1-p)\theta_j$. If $p \approx 1$, then selection based on the Y^*s will be close to selection based on the Ys, but the $Y^\dagger$s will be small, so there will be little power for inference, and conversely if $p \approx 0$.

4. The *Simes procedure* for simultaneously testing null hypotheses $H_1, \ldots, H_m$ with respective p-values $P_1, \ldots, P_m$ ordered as $P_{(1)} \le \cdots \le P_{(m)}$ rejects the global null hypothesis $H_0 = H_1 \cap \cdots \cap H_m$ if $P_{(k)} \le k\alpha/m$ for at least one k.

   (a) If $U_{(1)} \le \cdots \le U_{(n)}$ are the order statistics of $n > 1$ independent $U(0,1)$ variables, show that the conditional distribution of $U_{(1)}/u_n \le \cdots \le U_{(n-1)}/u_n$ given that $U_{(n)} = u_n$ is that of the order statistics of $n-1$ independent $U(0,1)$ variables. **Solution**: If $X_1, \ldots, X_n i.i.d. f$ are continuous random variables then the joint density of the corresponding order statistics $X_{(1)} \le \cdots \le X_{(n)}$ is

   $$f_{X_{(1)},\ldots,X_{(n)}}(x_1, \ldots, x_n) = n! f(x_1) \cdots f(x_n) I(x_1 \le \cdots \le x_n),$$

   so the joint density of $U_{(1)} \le \cdots \le U_{(n)}$ is

   $$f_{U_{(1)},\ldots,U_{(n)}}(u_1, \ldots, u_n) = n! I(0 \le u_1 \le \cdots \le u_n \le 1).$$

   Moreover

   $$\Pr(U_{(n)} \le u_n) = \Pr(U_1 \le u_n, \ldots, U_n \le u_n) = \prod_{j=1}^{n} \Pr(U_j \le u_n) = u_n^n, \quad 0 \le u_n \le 1,$$

   so the density of $U_{(n)}$ is $n u_n^{n-1}$, for $0 \le u_n \le 1$. Hence the joint conditional density is

   $$\begin{aligned} f_{U_{(1)},\ldots,U_{(n-1)} \mid U_{(n)}}(u_1, \ldots, u_{n-1} \mid u_n) &= \frac{f_{U_{(1)},\ldots,U_{(n)}}(u_1, \ldots, u_n)}{f_{U_{(n)}}(u_n)} \\ &= \frac{n!}{n u_n^{n-1}} I(0 \le u_1 \le \cdots \le u_{n-1} \le u_n \le 1), \end{aligned}$$

   and the change of variables $U'_{(1)} = U_{(1)}/u_n, \ldots, U'_{(n-1)} = U_{(n-1)}/u_n$ yields

   $$f_{U'_{(1)},\ldots,U'_{(n-1)} \mid U_{(n)}}(u'_1, \ldots, u'_{n-1} \mid u_n) = (n-1)! I(u'_1 \le \cdots \le u'_{n-1} \le 1),$$

   which is of the same form as the joint density of $U_{(1)} \le \cdots \le U_{(n)}$ but with $n-1$ instead of n.

   (b) Let $\alpha \in [0,1]$ and define $A_m(\alpha) = \Pr(P_{(k)} > k\alpha/m, k = 1, \ldots, m)$. Show by induction that under H_0 and when the P_k are independent, $A_m(\alpha) = 1 - \alpha$ for any m, and deduce that the familywise error rate of the Simes procedure is exactly α. For the induction, condition on the value of $P_{(m+1)}$. **Solution**: As $P_{(1)} = P_1$ for a sample of size $m = 1$ and $P_1 \sim U(0,1)$, we have

   $$A_1(\alpha) = \Pr(P_{(1)} > \alpha) = \Pr(P_1 > \alpha) = 1 - \alpha,$$

i.e., the result is true for $m = 1$. To establish the induction, suppose it is true for some $m \geq 1$. Then

$$\begin{aligned}
A_{m+1}(\alpha) &= \Pr\{P_{(k)} > k\alpha/(m+1), k = 1, \ldots, m+1\} \\
&= \int_\alpha^1 \Pr\{P_{(k)} > k\alpha/(m+1), k = 1, \ldots, m \mid P_{(m+1)} = u\}\,(m+1)u^m \,\mathrm{d}u \\
&= \int_\alpha^1 \Pr\left(P'_{(k)} > (k/m)[\alpha m/\{(m+1)u], k = 1, \ldots, m\right)\,(m+1)u^m \,\mathrm{d}u \\
&= \int_\alpha^1 A_m\,[\alpha m/\{(m+1)u\}]\,(m+1)u^m \,\mathrm{d}u \\
&= \int_\alpha^1 [1 - \alpha m/\{(m+1)u\}]\,(m+1)u^m \,\mathrm{d}u \\
&= \int_\alpha^1 \left\{(m+1)u^m - \alpha m u^{m-1}\right\} \,\mathrm{d}u \\
&= \left[u^{m+1} - \alpha u^m\right]_\alpha^1 = 1 - \alpha,
\end{aligned}$$

where the $P'_{(k)} = P_{(k)}/u$ are the order statistics of a uniform sample of size m, using part (a). This establishes the induction and gives FWER $= 1 - A_m(\alpha) = \alpha$ for any m.

(c) Which of the Simes and Bonferroni procedures is preferable? **Solution**: Both procedures have FWER less than or equal to α and respective rejection regions $\mathcal{Y}_B = \{P_{(m)} \leq \alpha/m\}$ and $\mathcal{Y}_S = \{P_{(m)} \leq \alpha/m, \ldots, P_{(1)} \leq \alpha\}$. As $\mathcal{Y}_B \subset \mathcal{Y}_S$ the Simes procedure must be more powerful (there are more ways to reject H_0). Hence the Simes procedure should always be preferred when the p-values are independent. The proof in (b) shows that it has exact FWER α, whereas the Bonferroni FWER is less than or equal to α, and this results in a loss of power. On the other hand the Bonferroni argument works also when the tests are dependent.