HW6, Theory of Inference 2016/7

Jonathan Rougier
School of Mathematics
University of Bristol UK

In the following questions, I show marks in square brackets, to give you an idea of the approximate tariff the question would carry in an exam. In these questions, the model is the usual $\{\mathcal{Y}, \Omega, f\}$.

1. Define a 'family of confidence procedures'. What additional property do you think that 'nesting' families of confidence procedures have? [5 marks]

   Answer. The function $C : \mathcal{Y} \times [0, 1] \to 2^{\Omega}$ is a family of confidence procedures exactly when $C(\bullet; \alpha)$ is a level $1 - \alpha$ confidence procedure for every $\alpha \in [0, 1]$.

   A nesting family would have $C(y, \alpha) \subset C(y; \alpha')$ for all $y \in \mathcal{Y}$ whenever $\alpha' \leq \alpha$. I.e., a confidence set with any specific level contains the confidence sets of all lower levels.

2. Define a 'family of significance procedures'. Prove that if p is a family of significance procedures, then $\sup_{\theta \in \Omega_0} p(y; \theta)$ is a significance procedure for the null hypothesis $\Omega_0 \subset \Omega$. [10 marks]

   Answer. The function $p : \mathcal{Y} \times \Omega \to \mathbb{R}$ is a family of significance procedures exactly when $p(\bullet; \theta)$ is a significance procedure for the null model $\Omega_0 = \{\theta\}$ for every $\theta \in \Omega$. That is, $p(Y; \theta)$ is super-uniform under $Y \sim f(\bullet; \theta)$ for every $\theta \in \Omega$.

   We must show that $p(y; \Omega_0) := \sup_{\theta \in \Omega_0} p(y; \theta)$ is super-uniform for every $\theta \in \Omega_0$. Now

   $$p(y; \Omega_0) \leq u \implies p(y; \theta) \leq u \quad \text{for all } \theta \in \Omega_0.$$

   Hence, by the probability inequality for implication,

   $$\begin{aligned} \Pr\{p(Y; \Omega_0) \leq u; \theta\} &\leq \Pr\{p(Y; \theta) \leq u; \theta\} && \text{for all } \theta \in \Omega_0 \\ &\leq u && p \text{ is a family of confidence procedures} \end{aligned}$$

showing that $p(Y; \Omega_0)$ is super-uniform for all $\theta \in \Omega_0$.

3.

Theorem 1 (The Duality Theorem)**.** *Let p be a family of signficance procedures. Then*

$$C(y; \alpha) = \big\{\theta \in \Omega : p(y; \theta) > \alpha\big\}$$

is a nesting family of confidence procedures. Conversely, let C be a nesting family of confidence procedures. Then

$$p(y; \theta) = \inf \big\{\alpha : \theta \notin C(y; \alpha)\big\}$$

is a family of significance procedures.

Prove the first half of this theorem. [10 marks]

Answer. We must show that if p is a family of significance procedures, then $C(\bullet; \alpha)$ is a level $1 - \alpha$ confidence procedure. So, to compute the coverage of C:

$$\begin{aligned}
\Pr\{\theta \in C(Y; \alpha); \theta\} &= \Pr\{p(Y; \theta) > \alpha; \theta\} \\
&= 1 - \Pr\{p(Y; \theta) \le \alpha; \theta\} \\
&= 1 - (\alpha - \epsilon) && \text{for some } \epsilon \ge 0 \text{, because } p(Y; \theta) \text{ is super-uniform} \\
&= 1 - \alpha + \epsilon \ge 1 - \alpha
\end{aligned}$$

as needed to be shown.

4. We proved in the lecture that

$$p(y; \theta) = \Pr\{g(Y) \ge g(y); \theta\}$$

is a family of significance procedures, for any $g : \mathcal{Y} \to \mathbb{R}$. Prove this result directly in the special case of $g(y) = c$, where c is any constant. [10 marks]

Answer. Let $u \in [0, 1]$ be arbitrary. Then

$$\begin{aligned}
\Pr\{p(Y; \theta) \le u; \theta\} &= \Pr\big\{\Pr\{c \ge c; \theta\} \le u; \theta\big\} \\
&= \Pr\{1 \le u; \theta\} \\
&= \begin{cases} 0 & 0 \le u < 1 \\ 1 & u = 1. \end{cases}
\end{aligned}$$

In both branches we have a value less than or equal to u, confirming the theorem.

5. If your p-value is small, then the observations are improbable under your null hypothesis. What information do you need to compute the probability that your null hypothesis is true, given the observations? [10 marks]

Answer. We can think about this using Bayes's theorem:

$$\Pr(H_0 \mid Y = y) = \frac{\Pr(Y = y \mid H_0) \cdot \Pr(H_0)}{\Pr(Y = y)}$$

If we replace the observables with a significance procedure and a small value u then we get

$$\begin{aligned}\Pr\{H_0 \mid p(Y; H_0) \leq u\} &= \frac{\Pr\{p(Y; H_0) \leq u \mid H_0\} \cdot \Pr(H_0)}{\Pr(Y = y)} \\ &\leq \frac{u \cdot \Pr(H_0)}{\Pr\{p(Y; H_0) \leq u\}}.\end{aligned}$$

So we need to know $\Pr(H_0)$ and $\Pr\{p(Y; H_0) \leq u\}$. This latter term can be expanded as (Law of Total Probability)

$$\Pr\{p(Y; H_0) \leq u\} = \Pr\{p(Y; H_0) \leq u | H_0\} \cdot \Pr(H_0) + \Pr\{p(Y; H_0) \leq u | \neg H_0\} \cdot \{1 - \Pr(H_0)\}$$

where $\neg H_0 = \Omega \setminus H_0$. So as well as $\Pr(H_0)$ we also need to know the probability distribution of the p-value for H_0 under the complement of H_0 in Ω.

The point is, computing a p-value does not require us to know these quantities, which we would need to know to assess the probability that H_0 were true; hence a p-value cannot tell us about this.