HW8

Sam Olson

Q1

Suppose there is one observation X with pdf

$$f(x) = 2\theta(1-2x) + 2x, \quad \text{for } x \in [0,1],\, \theta \in [0,1].$$

Find the Bayes test for

$$H_0 : \theta \leq 0.4 \quad \text{vs.} \quad H_1 : \theta > 0.4$$

with respect to the uniform prior on $[0,1]$.

Answer

We are given the likelihood:

$$f(x \mid \theta) = 2\theta(1-2x) + 2x, \quad x \in [0,1], \theta \in [0,1]$$

and a uniform prior:

$$\pi(\theta) = 1, \quad \text{for } \theta \in [0,1]$$

We compute the (unnormalized) posterior:

$$\pi(\theta \mid x) \propto f(x \mid \theta)\pi(\theta) = 2\theta(1-2x) + 2x$$

Normalize:

$$\int_0^1 [2\theta(1-2x) + 2x]\, d\theta = (1-2x) \cdot 1 + 2x \cdot 1 = 1$$

so:

$$\pi(\theta \mid x) = 2\theta(1-2x) + 2x$$

Let us define the posterior probabilities of the hypotheses:

$$P(H_0 \mid x) = \int_0^{0.4} \pi(\theta \mid x)\, d\theta = (1-2x) \cdot (0.4)^2 + 2x \cdot 0.4 = 0.16(1-2x) + 0.8x$$

Simplifying:

$$P(H_0 \mid x) = 0.16 + 0.48x$$

$$P(H_1 \mid x) = 1 - P(H_0 \mid x) = 0.84 - 0.48x$$

The Bayes test rejects H_0 when $P(H_1 \mid x) > P(H_0 \mid x)$, that is:

$$0.84 - 0.48x > 0.16 + 0.48x \Rightarrow 0.68 > 0.96x \Rightarrow x < \frac{17}{24}$$

We define the (Bayes) test function $\varphi(x)$ as:

$$\varphi(x) = \begin{cases} 1, & \text{if } x < \dfrac{17}{24}, \\ \\ 0, & \text{if } x \geq \dfrac{17}{24} \end{cases}$$

That is, we reject H_0 if $x < \frac{17}{24}$, and fail to reject H_0 otherwise.

Q2

Problem 9.13, Casella and Berger (2nd Edition)

Let X be a single observation from the $\text{Beta}(\theta, 1)$ pdf.

a)

Let $Y = -(\log X)^{-1}$. Evaluate the confidence coefficient of the set $[y/2, y]$.

Answer

If $X \sim \text{Beta}(\theta, 1)$, then its pdf is:

$$f_X(x) = \theta x^{\theta-1}, \quad 0 < x < 1$$

Define the transformation:

$$Y = -\frac{1}{\log X} \Rightarrow X = e^{-1/Y}, \quad Y > 0$$

Compute the pdf of Y via the change of variables:

$$f_Y(y) = f_X(e^{-1/y}) \cdot \left| \frac{d}{dy} e^{-1/y} \right| = \theta \cdot e^{-\theta/y} \cdot \frac{1}{y^2}, \quad y > 0$$

So the pdf of Y is:

$$f_Y(y) = \frac{\theta}{y^2} e^{-\theta/y}, \quad y > 0$$

We seek:

$$P\left(\frac{Y}{2} \le \theta \le Y\right) = P\left(\theta \in \left[\frac{Y}{2}, Y\right]\right)$$

This is equivalent to:

$$P\left(\theta \in \left[\frac{Y}{2}, Y\right]\right) = P\left(Y \in [\theta, 2\theta]\right)$$

Thus:

$$\text{Confidence Coefficient} = \int_{\theta}^{2\theta} \frac{\theta}{y^2} e^{-\theta/y}\, dy$$

Make substitution $u = \theta/y \Rightarrow y = \theta/u$, $dy = -\theta/u^2 du$:

$$\text{Confidence Coefficient} = \int_{1/2}^{1} e^{-u}\, du = e^{-1/2} - e^{-1} \approx 0.6065 - 0.3679 = 0.2386$$

b)

Find a pivotal quantity and use it to set up a confidence interval having the same confidence coefficient as part a).

Answer

The pdf of $X \sim \text{Beta}(\theta, 1)$ is:

$$f_X(x) = \theta x^{\theta-1}, \quad x \in (0,1)$$

Let us consider the transformation:

$$T = X^\theta$$

Then the cdf of T is:

$$P(X^\theta \leq t) = P(X \leq t^{1/\theta}) = \int_0^{t^{1/\theta}} \theta x^{\theta-1} dx = t$$

Thus, $T = X^\theta \sim \text{Uniform}(0,1)$, and is a pivotal quantity.

We want to find values $a, b \in (0,1)$ such that:

$$P(a \leq T \leq b) = b - a = 0.239$$

This implies:

$$P(a \leq X^\theta \leq b) = 0.239$$

Solving for θ from $a \leq X^\theta \leq b$:

$$a \leq X^\theta \leq b \Rightarrow \frac{\log a}{\log X} \leq \theta \leq \frac{\log b}{\log X}, \quad \text{(since } 0 < X < 1 \text{ and } \log X < 0)$$

c)

Compare the two confidence intervals.

Answer

The interval in part a), $[Y/2, Y]$, depends on the transformed variable $Y = -1/\log X$, and is symmetric on the log scale.

The interval in part b) uses a pivotal quantity to define a confidence set for θ.

The part a) interval is a special case of the pivotal-based interval in part b), using fixed endpoints $a = e^{-1}, b = e^{-1/2}$ such that $b - a = 0.239$.

The pivotal method in part b) allows for more flexibility and can produce shorter intervals.

For example, choosing $b = 1$, $a = 1 - 0.239$, yields:

$$\theta \in \left[0, \frac{\log(1-0.239)}{\log X}\right]$$

This is a one-sided interval of the same confidence level, and is often shorter or more desirable depending on the application.

Q3

Problem 9.16, Casella and Berger (2nd Edition)

Let $X_1, \dots, X_n$ be i.i.d. $\mathrm{N}(\theta, \sigma^2)$, where σ^2 is known. For each of the following hypotheses, write out the acceptance region of a level α test and the $1-\alpha$ confidence interval that results from inverting the test.

a)

$H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$

Answer

$$Z = \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1) \quad \text{under } H_0$$

Reject H_0 if:

$$|Z| > z_{\alpha/2} \quad \Longleftrightarrow \quad \left|\bar{X} - \theta_0\right| > z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Accept H_0 if:

$$\theta_0 \in \left[\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

Inverting the test yields the $(1-\alpha)$ confidence interval for θ:

$$\left[\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

b)

$H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$

Answer

$$Z = \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}}$$

Reject H_0 if:

$$Z < -z_\alpha \quad \Longleftrightarrow \quad \bar{X} - \theta_0 < -z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

Accept H_0 if:

$$\bar{X} \geq \theta_0 - z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

Inverting the test yields the one-sided confidence interval:

$$\left(-\infty,\, \bar{X} + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}\right]$$

c)

$H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$

Answer

$$Z = \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}}$$

Reject H_0 if:

$$Z > z_\alpha \quad \iff \quad \bar{X} - \theta_0 > z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

Accept H_0 if:

$$\bar{X} \leq \theta_0 + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

Inverting the test gives the one-sided interval:

$$\left[\bar{X} - z_\alpha \cdot \frac{\sigma}{\sqrt{n}}, \infty\right)$$

Q4

Problem 9.11, Casella and Berger (2nd Edition)

If T is a continuous random variable with cdf $F_T(t \mid \theta)$ and $\alpha_1 + \alpha_2 = \alpha$, show that an α-level acceptance region of the hypothesis $H_0 : \theta = \theta_0$ is

$$\{t : \alpha_1 \leq F_T(t \mid \theta_0) \leq 1 - \alpha_2\},$$

with associated confidence $1 - \alpha$ set

$$\{\theta : \alpha_1 \leq F_T(t \mid \theta) \leq 1 - \alpha_2\}.$$

Answer

Let T be a continuous test statistic with cumulative distribution function $F_T(t \mid \theta)$.

Under the null hypothesis $H_0 : \theta = \theta_0$, we define the transformed variable:

$$U = F_T(T \mid \theta_0)$$

Since T is continuous and $F_T(\cdot \mid \theta_0)$ is strictly increasing, the probability integral transform implies:

$$U \sim \text{Uniform}(0, 1) \quad \text{under } H_0$$

We construct an acceptance region that excludes the lower α_1 and upper α_2 tails of this uniform distribution. Specifically, we accept H_0 if:

$$\alpha_1 \leq F_T(t \mid \theta_0) \leq 1 - \alpha_2$$

The probability of rejecting H_0 is the sum of the tail probabilities:

$$P_{\theta_0}\left(F_T(T \mid \theta_0) < \alpha_1\right) = \alpha_1$$

And:

$$P_{\theta_0}\left(F_T(T \mid \theta_0) > 1 - \alpha_2\right) = \alpha_2$$

Hence, the total probability of rejection under H_0 is:

$$P_{\theta_0}(\text{Reject } H_0) = \alpha_1 + \alpha_2 = \alpha$$

So this acceptance region defines a level α test.

We now construct a $1 - \alpha$ confidence set by inverting the acceptance region:

Fix observed value t_{obs}, and define the set of all parameter values θ for which t_{obs} lies within the corresponding acceptance region:

$$C(t_{\text{obs}}) = \{\theta : \alpha_1 \leq F_T(t_{\text{obs}} \mid \theta) \leq 1 - \alpha_2\}$$

This set contains all values of θ that are not rejected when testing $H_0 : \theta = \theta'$ for each possible θ'.

Since $F_T(T \mid \theta) \sim \text{Uniform}(0, 1)$ under the true parameter θ, we have:

$$P_\theta \left(\alpha_1 \leq F_T(T \mid \theta) \leq 1 - \alpha_2 \right) = 1 - \alpha_1 - \alpha_2 = 1 - \alpha$$

Thus, the random set

$$\{\theta : \alpha_1 \leq F_T(t \mid \theta) \leq 1 - \alpha_2\}$$

is a confidence set for θ with coverage probability $1 - \alpha$.

Acceptance region (level α):

$$\{t : \alpha_1 \leq F_T(t \mid \theta_0) \leq 1 - \alpha_2\}$$

Associated confidence set (level $1 - \alpha$):

$$\{\theta : \alpha_1 \leq F_T(t \mid \theta) \leq 1 - \alpha_2\}$$