ORFE 524: Statistical Theory and Methods Homework 4

Zachary Hervieux-Moore

Monday 21^{st} November, 2016

Exercise 1: Let ϕ denote the c.d.f. of $N(0,1)$, and let $z_\alpha \in \mathbb{R}$ denote the critical value satisfying $\phi(z_\alpha) = 1-\alpha$. Consider the following regression with fixed design:

$$y = X\beta + \eta$$

where $X \in \mathbb{R}^{n\times d}$ is a nonrandom matrix, $\beta \in \mathbb{R}^d$, and $\eta \sim N(0, I_n)$. Suppose that X^TX has full rank, so that $\widehat{\beta} = (X^TX)^{-1}X^Ty$ is well defined.

1) Derive an ellipsoidal $(1-\alpha)$-confidence set for $\beta \in \mathbb{R}^d$ in terms of z_α. Remember that $\{X : \|Ax\| \leq r\}$, $A \succeq 0$ is an ellipsoid.

2) Derive a hypercubic $(1-\alpha)$-confidence set for $\beta \in \mathbb{R}^d$, which is of the form $\{\beta : \|\beta - c\|_\infty < r\}$. Assume now that $\eta \sim N(0, \sigma^2 I_n)$, where $\sigma^2 > 0$ is unknown.

3) Suppose now that the design matrix X is also random. Are the confidence sets in 1) and 2) still of $1-\alpha$?

Answer:

1) We first note that

$$\widehat{\beta} \sim N(\beta, (X^TX)^{-1})$$

Which means that

$$\begin{gathered}\widehat{\beta} - \beta \sim N(0, X^TX)^{-1}) \\ (X^TX)^{1/2}(\widehat{\beta} - \beta) \sim N(0, I_d) \\ (\widehat{\beta} - \beta)^T(X^TX)(\widehat{\beta} - \beta) \sim N(0, I)_d^2\end{gathered}$$

Noting that $(\widehat{\beta} - \beta)^T(X^TX)(\widehat{\beta} - \beta) = \|(X^TX)^{1/2}(\widehat{\beta} - \beta)\|^2$. Hence, the confidence set becomes

$$\{\beta : \|(X^TX)^{1/2}(\widehat{\beta} - \beta)\| \leq z_{\alpha/2}\}$$

2) Now that σ is unknown, we have

$$\widehat{\beta} \sim N(\beta, (X^TX)^{-1}\sigma^2)$$

Thus, we have for all j

$$\frac{\widehat{\beta}_j - \beta_j}{\widehat{\sigma}\sqrt{v_j}} \sim N(0,1)$$

where v_j is the j^{th} entry on the diagonal of $(X^T X)^{-1}$.

Now, $\|\cdot\|_\infty < r$ implies we must have all entries less than r. Thus to ensure that all entries have a probability less than r, we must have each individual probability less than $r^{1/d}$. Thus the confidence set is

$$\{\beta : \|\widehat{\beta} - \beta\|_\infty < z_{\alpha/2}^{1/d} \cdot \widehat{\sigma} \cdot \sqrt{v}\}$$

Where all the inequality is element-wise.

3) Yes, the confidence set are still $1 - \alpha$. Since we conditioned on X in the previous parts, if we further condition that X is random, by the tower property, it simply returns the equivalent probabilities.

Exercise 2: Suppose that in 1), β is k-sparse, i.e., having exactly k nonzero coordinates $i_1, \ldots, i_k$. We further assume that $X^T X = nI_d$. Design a procedure that identifies $\text{supp}(\beta) = \{i_1, \ldots, i_k\}$. How large should n be for the procedure to be successful with probability greater than $1-\alpha$? Try to find an n as small as possible, and express that quantity in terms of $\kappa := \min_k |\beta_{i_k}|$.

Answer: The procedure we will use is to sort the β_i's by absolute value and pick the k largest. From question 1), we have

$$\widehat{\beta} \sim N(\beta, (X^T X)^{-1}) = N(\beta, \frac{1}{n} I_d)$$

In particular,

$$\widehat{\beta}_j \sim N(\beta_j, \frac{1}{n})$$
$$\sqrt{n}(\widehat{\beta}_j - \beta_j) \sim N(0,1)$$

Thus, we want all the $\beta_j = 0$ such that

$$P(|\beta_j| < |\beta_{i_k}|) = 1 - \alpha$$

Which we know is the confidence set for the zero β_j's are

$$\widehat{\beta}_j \in 0 \pm z_{\alpha/2}/n$$

Thus we pick $z_{\alpha/2}/n < \kappa$ or pick n sufficiently large such that $z_{\alpha/2} < \kappa n$.

Exercise 3:

1) Let $x \sim \text{Uniform}^n[0, \theta]$. Find a $(1-\alpha)$-confidence set for θ.

2) Suppose that each $P \in \mathcal{P}$ has median $m = m(P)$, $x \sim P^n$. What is the confidence coefficient of the intervals $[x_{(1)}, \infty)$, $(-\infty, x_{(n)}]$, and $[x_{(1)}, x_{(n)}]$?

Answer:

1) Recall that $x_{(n)}$ has pdf

$$P(x_{(n)} \leq x) = \left(\frac{x}{\theta}\right)^n$$

Then we note the following

$$P(x_{(n)} \leq \theta) = 1$$

$$P(x_{(n)} \geq \alpha^{1/n}\theta) = 1 - P(x_{(n)} \leq \alpha^{1/n}\theta) = 1 - \left(\frac{\alpha^{1/n}\theta}{\theta}\right)^n = 1 - \alpha$$

Thus, the $(1-\alpha)$-confidence interval is $\theta \in [x_{(n)}, x_{(n)}/\alpha^{1/n}]$

2) We define a new random variable $Y = \#$ of $X_i < m$. That is, the number of samples less than the median. Y is distributed as Binom(n, 1/2) because each X_i has a 1/2 probability being below m since it is the median. Also, each sample is independent, thus it becomes a sum of Bernoulli trials. Then,

$$\begin{aligned} P(X_{(1)} < m < \infty) &= \sum_{i=1}^{\infty} P(Y = i) \\ &= \sum_{i=1}^{n} P(Y = i) = \sum_{i=1}^{n} \binom{n}{i} \frac{1}{2^n} \\ &= 1 - \binom{n}{0} \frac{1}{2^n} = 1 - \frac{1}{2^n} \end{aligned}$$

Following the same logic,

$$\begin{aligned} P(-\infty < m < X_{(n)}) &= 1 - P(X_{(n)} \leq m) \\ &= 1 - P(Y = n) = 1 - \binom{n}{n} \frac{1}{2^n} \\ &= 1 - \frac{1}{2^n} \end{aligned}$$

Finally,

$$\begin{aligned}
P(X_{(1)} < m < X_{(n)}) &= P(1 < Y < n) \\
&= \sum_{i=1}^{n-1} \binom{n}{i} \frac{1}{2^n} \\
&= 1 - \frac{1}{2^n} - \frac{1}{2^n} \\
&= 1 - \frac{1}{2^{n-1}}
\end{aligned}$$

Exercise 4: Suppose that we have observed $x := \{x_i\} \sim N^n(\mu, \sigma^2)$, where σ^2 is known. Consider the hypothesis test: $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$. We know that, $T(x) = 1$ (rejecting the null hypothesis) iff $|\sqrt{n}(\bar{x} - \mu_0)/\sigma| > z_{\alpha/2}$ has size α. Let $\beta_T(\mu)$ be the power function. Derive $\beta_T(\mu)$ in terms of ϕ, the standard normal cdf. Draw an outline of this function. How does $B_T(\mu)$ behave as $n \to \infty$?

Answer: Recall the definition of the power of a test,

$$P(T(x) = 1|\mu \neq \mu_0)$$
$$P(|\sqrt{n}(\bar{x} - \mu_0)/\sigma| > z_{\alpha/2}|\mu \neq \mu_0)$$
$$= 1 - P(-z_{\alpha/2} < \sqrt{n}(\bar{x} - \mu_0 + \mu - \mu)/\sigma < z_{\alpha/2}|\mu \neq \mu_0)$$
$$= 1 - P(-z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}/\sigma < \sqrt{n}(\bar{x} - \mu)/\sigma < z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}/\sigma|\mu \neq \mu_0)$$

But since $\mu \neq \mu_0$, then $\sqrt{n}(\bar{x} - \mu)/\sigma \sim N(0, 1)$. Thus,

$$= 1 - \phi(z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}/\sigma) + \phi(-z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}/\sigma)$$

As $n \to \infty$, both cdf's will tend to 1 and thus the power will tend to 1 as well. Also note that at as $\mu \to \mu_0$, we get that the power tends to α.

Exercise 5: Let $x = \{x_i\} \sim \text{Bern}^n(p)$, where $0 < p < 1$. Consider the following hypothesis testing problem: $H_0 : p \geq p_0$, versus $H_1 : p < p_0$.

1) Construct an α-level test T_α depending on $\bar{x}$.

2) Derive another α-level test using the p-value for $\{T_\alpha\}$.

Answer:

We shall use Hoeffding's inequality which states for bounded random variables,

$$P(|\bar{x} - p| \geq t) \leq 2e^{-2nt^2}$$

Thus pick $t = \sqrt{-\frac{1}{2n}\log(\alpha/2)}$

$$P(|\bar{x} - p| \geq t) \leq \alpha$$
$$P(|\bar{x} - p| \leq t) \geq 1 - \alpha$$

This is a two-tailed test, but we only need one.

Thus construct the test

$$T_\alpha = 1_{\{\bar{x} > p_0 + \sqrt{-\frac{1}{2n}\log(\alpha)}\}}$$

Which has level α.

We just use the definition of the p-value.

$$\inf_\alpha P(x > t_\alpha) \leq \alpha$$

And define the test

$$T_\alpha = 1_{\{x > t_\alpha\}}$$

Exercise 6:

1) Suppose that T_{θ_0} is an α-size for any $H_0 : \theta = \theta_0$ versus $H_1 : \theta \in \Theta_1$ where $\theta_0 \notin \Theta_1$. Let

$$S(x) = \{\theta : T_\theta(x) = 0, \text{i.e. } x \in \mathcal{A}(T_\theta)\}$$

Is it true that the confidence coefficient for S is $1 - \alpha$? Argue for your answer.

2) Let $x \sim N^n(\mu, \sigma^2)$, where μ and σ^2 are unknown. Obtain a $(1 - \alpha)$ confidence set by inverting some T.

Answer:

1) We have since $x \in \mathcal{A}(T_\theta) \implies \theta \in S(X)$,

$$P(\theta \in S(X)) \geq P(x \in \mathcal{A}(T_\theta)) = 1 - P(x \in \mathcal{R}(T_\theta)) = 1 - \alpha$$

Since T_{θ_0} is an α-size means $P(x \in \mathcal{R}(T_\theta)) = \alpha$. Thus,

$$P(\theta \in S(X)) \geq 1 - \alpha$$

And so $S(X)$ is a $(1 - \alpha)$confidence set.

2) Define the test

$$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$$

Since we have σ^2 unknown, we have

$$\frac{\bar{x} - \mu}{S_{n-1}/\sqrt{n}} \sim \text{t-distribution}(n-1)$$

Thus, the test becomes,

$$T = 1_{\{|\bar{x}-\mu| \geq t_{\alpha/2} S_{n-1}/\sqrt{n}\}}$$

Which has probability of rejection of α. So,

$$P(\mathcal{A}(T)) = 1 - \alpha = P(|\bar{x} - \mu| \leq t_{\alpha/2} S_{n-1}/\sqrt{n})$$

Thus the confidence set is

$$\mu \in \bar{x} \pm t_{\alpha/2} S_{n-1}/\sqrt{n}$$

Exercise 7: Consider a hypothesis test $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$ for $X \sim P_\theta \in \mathcal{P}$. Let

$$T(x) = 1_{\{\tau(x) \leq \tau_0\}}$$

be a test function with size α, where τ is sufficient and $\tau_0 \in \mathbb{R}$. Suppose X and τ have densities f_X^θ and f_τ^θ with respect to some measure σ for all θ.

1) Consider a new hypothesis test $H_0' : \theta = \theta_0$ versus $H_1' : \theta = \theta_1$. Give a simple argument why an α-size test

$$T(x) = 1_{\{f_\tau^{\theta_0}(\tau(x)) \leq C \cdot f_\tau^{\theta_1}(\tau(x))\}}$$

is UMP, where C is a constant.

2) Assume X and τ are continuous random variables. Derive and argue sufficient conditions on $f_\tau^\theta / f_\tau^{\theta'}$ to ensure the test function T defined in above is UMP.

3) Give a concrete example where your conditions hold.

Answer:

1) Let T'', $P_{\theta_0}(R(T'')) \leq \alpha = P_{\theta_0}(R(T'))$. Which implies,

$$P_{\theta_0}(R(T'')\backslash R(T')) \leq P_{\theta_0}(R(T')\backslash R(T'))$$

Now we show $P_{\theta_1}(R(T')\backslash R(T')) \leq P_{\theta_1}(R(T')\backslash R(T''))$

$$P_{\theta_1}(R(T')\backslash R(T')) \leq \frac{1}{c} P_{\theta_0}(R(T')\backslash R(T'))$$
$$\leq \frac{1}{c} P_{\theta_0}(R(T')\backslash R(T')) \leq \frac{c}{c} P_{\theta_1}(R(T')\backslash R(T''))$$

Thus,

$$P_{\theta_1}(R(T'')\backslash R(T')) + P(R(T') \cap R(T''))$$
$$\leq P_{\theta_1}(R(T')\backslash R(T'')) + P(R(T') \cap R(T''))$$
$$\implies P_{\theta_1}(R(T'')) \leq P_{\theta_1}(R(T'))$$

Thus, T' is UMP.

2) This is a version of Karl-Rubin and so we require for $\theta > \theta'$, $\phi_{\theta,\theta'} = \frac{f_\theta(x)}{f_{\theta'}(x)}$ is increasing on $\mathbb{R}$.

3) A concrete example is $x = \{x_i\} \sim N^n(\mu, \sigma^2)$, known σ^2, $\bar{x} \sim N(\mu, \sigma^2/n)$. Here, $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$. If $\mu > \mu'$, then dividing the Gaussians will yield a function that is increasing.

Exercise 8:

1) Let $\{x_i\} \sim N^n(\mu_1, 1)$, and $\{y_i\} \sim N^n(\mu_2, 1)$. Suppose that $\{x_i\}$ and $\{y_i\}$ are independent. Derive a α-size test for $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$.

2) Let $\{x_i\} \sim N(\mu, 1)$, and consider the testing problem $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. Consider α-size tests of the form $T_>(x) = 1_{\{\bar{x} \geq t\}}$, $T_>(x) = 1_{\{\bar{x} \leq t\}}$, and $T(x) = 1_{\{|\bar{x}-\mu_0| \geq t\}}$, each with an appropriate value of t. Discuss their relative power and order them by power (over regions of the real line).

Answer:

1) First note that we can rewrite the test as $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \neq 0$. Thus defining $z_i = x_i - y_i \sim N^n(\mu_1 - \mu_2, 1)$. We obtain that this test is the typical test for Gaussians. Thus, the α-size test is

$$T_\alpha(z) = 1_{\{|\sqrt{n}(\bar{x}-(\mu_1-\mu_2))| > z_\alpha\}}$$

2) Computing the power of $T_>$

$$\begin{aligned} &P(T_>(x) = 1 | \mu \neq \mu_0) \\ &= P(\sqrt{n}(\bar{x} - \mu_0) > z_\alpha | \mu \neq \mu_0) \\ &= 1 - P(\sqrt{n}(\bar{x} - \mu_0 + \mu - \mu) < z_\alpha | \mu \neq \mu_0) \\ &= 1 - P(\sqrt{n}(\bar{x} - \mu) < z_\alpha + (\mu_0 - \mu)\sqrt{n} | \mu \neq \mu_0) \\ &= 1 - \phi(z_\alpha + (\mu_0 - \mu)\sqrt{n}) \end{aligned}$$

Likewise for $T_<$

$$\begin{aligned} &P(T_<(x) = 1 | \mu \neq \mu_0) \\ &= P(\sqrt{n}(\bar{x} - \mu_0) < -z_\alpha | \mu \neq \mu_0) \\ &= P(\sqrt{n}(\bar{x} - \mu_0 + \mu - \mu) < -z_\alpha | \mu \neq \mu_0) \\ &= P(\sqrt{n}(\bar{x} - \mu) < -z_\alpha + (\mu_0 - \mu)\sqrt{n} | \mu \neq \mu_0) \\ &= \phi(-z_\alpha + (\mu_0 - \mu)\sqrt{n}) \end{aligned}$$

Then using our result from question 4)

$$\begin{aligned} &P(T(x) = 1 | \mu \neq \mu_0) \\ &= 1 - \phi(z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}) + \phi(-z_{\alpha/2} + (\mu_0 - \mu)\sqrt{n}) \end{aligned}$$

Exercise 9: Let Y be a continuous r.v. with density f, and $E[Y] = \mu_Y$, $var(Y) - \sigma_Y^2$. A location-scale family based on f is

$$\mathcal{P} = \{P_{\mu,\sigma} \text{ with density } f_{\mu,\sigma}(x) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right), \sigma > 0, \mu \in \mathbb{R}\}$$

1) Let $x \sim P_{\mu,\sigma}$. Derive $E[X]$ and $var(X)$ in terms of μ_y and σ_Y^2.
2) Let $\{x_i\} \sim P_{\mu,\sigma}^n$, where σ is known. Argue for a pivotal quantity Z based on $\bar{x}$. Note that f is known.
3) Suppose $n = 2$. Derive the density f_Z of Z based on f. Derive a $(1-\alpha)$ confidence interval based on critical values of f_Z

Answer:

1) From the formula of transformation of random variables, it is evident that

   $$X = \sigma Y + \mu$$

   Thus, by linearity of expectation

   $$E[X] = \sigma\mu_Y + \mu$$

   As well,

   $$Var(X) = Var(\sigma Y + \mu) = \sigma^2 Var(Y) = \sigma^2\sigma_Y^2$$

2) We define $Z = \frac{X-\mu}{\sigma}$ which means, that $Z \sim Y$ thus Z has density f. Thus, $\bar{Z} = \frac{\bar{x}-\mu}{\sigma}$ is also a pivot variable since it is a sum of pivots.
3) We have that the confidence interval becomes

   $$\{\mu : \bar{x} - z_{\alpha/2}\sigma < \mu < \bar{x} + z_{\alpha/2}\sigma\}$$

   where z_α are the critical values for the pivot variable.

Exercise 10: Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$, where, for $0 < p < 1$, $X \sim pN_1(0, \sigma_1^2) + (1-p)N_2(\theta, \sigma_2^2)$. Assume we always know whether X is generated from N_1 or N_2. Let $l = 1$ or 2 denote which normal distribution x comes from. We also know σ_1 and σ_2. Suppose $\alpha > p$. Consider the following two tests:

- $T(x) = 1_{\{x > \theta_0 + z_\alpha \sigma_l\}}$, and
- $T'(x) = 1_{\{x > \theta_0 + z_{(\alpha-p)/(1-p)}\sigma_2, \text{ or } l=1\}}$

where z_α is the $(1-\alpha)$ quantile of a normal distribution. Show that these twi tests are of size α.

Answer: We first show that $T(x)$ is of size α.

$$\begin{gathered} P(T(x) = 1) = P(x > \theta_0 + z_\alpha \sigma_l) \\ = P(x > \theta_0 + z_\alpha \sigma_1) P(l = 1) + P(x > \theta_0 + z_\alpha \sigma_2) P(l = 2) \\ = \alpha p + \alpha(1-p) = \alpha \end{gathered}$$

Now for $T'(x)$,

$$\begin{gathered} P(T'(x) = 1) = P(\{x > \theta_0 + z_{(\alpha-p)/(1-p)}\sigma_2, \text{ or } l = 1\}) \\ = P(\{x > \theta_0 + z_{(\alpha-p)/(1-p)}\sigma_2, \text{ or } l = 1\}) P(l = 1) \\ +P(\{x > \theta_0 + z_{(\alpha-p)/(1-p)}\sigma_2, \text{ or } l = 1\}) P(l = 2) \\ = 1 \cdot p + (\alpha - p)/(1-p)(1-p) = p + \alpha - p = \alpha \end{gathered}$$