Statistical Theory MT 2009

Problems 2: Solution sketches

1. Denote the c.d.f. of a statistic $T = t(\mathbf{X})$ under the simple null hypothesis H_0 by F_{T,H_0}, and assume that F_{T,H_0} is continuous. Put

$$P(\mathbf{x}) = P(T \geq t(\mathbf{x})|H_0).$$

Let $\mathbf{X}$ be a random sample from the distribution specified by the null hypothesis. Show that then the random variable $P(\mathbf{X})$ is uniformly distributed on $[0, 1]$.

Solution: To derive this result, suppose that X has continuous distribution with invertible c.d.f. F, and $Y = F(X)$. Then $0 \leq Y \leq 1$, and for $0 \leq y \leq 1$,

$$P(Y \leq y) \;=\; P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y,$$

showing that $Y \sim \mathcal{U}[0, 1]$. It also follows that $1 - Y \sim \mathcal{U}[0, 1]$. Now note that

$$F_{T,H_0}(t) \;=\; P(T \leq t|H_0)$$

and so $F_{T,H_0}(T(\mathbf{X}))$ is of the form $F(X)$ for some random variable X.

2. Suppose you have a sample of size n from an exponential distribution with mean μ. Find the best size-α test of $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu = \mu_1$, where $\mu_1 > \mu_0$.

Solution: We have as density

$$f(x; \mu) = \frac{1}{\mu} e^{-\frac{x}{\mu}}, \quad x > 0,$$

giving as likelihood

$$L(\mu) = \mu^{-n} e^{-\frac{1}{\mu} \sum_{i=1}^n x_i}.$$

The Neyman-Pearson lemma tells us that the best test uses as test statistic the likelihood ratio

$$\frac{L(\mu_1)}{L(\mu_0)} = \left(\frac{\mu_0}{\mu_1}\right)^n e^{\sum_{i=1}^n x_i \left(\frac{1}{\mu_0} - \frac{1}{\mu_1}\right)}.$$

As $\mu_1 > \mu_0$ we have that $\frac{1}{\mu_0} - \frac{1}{\mu_1} > 0$, and thus we reject for large values of $\sum_{i=1}^n x_i$, or equivalently, for large values of $\overline{x}$.

We know that, under H_0, we have $\sum_{i=1}^n X_i \sim \Gamma(n, \mu_0^{-1})$, so the test statistic $T = \frac{1}{\mu_0} \sum_{i=1}^n X_i \sim \Gamma(n, 1)$, independent of μ, and so the critical region for T can be determined using the $\Gamma(n, 1)$-distribution.

Note: The critical region does not depend on the particular μ_1, so the test is UMP for $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$.

3. Suppose you have a random sample $\mathbf{X} = X_1, \ldots, X_n$ of size n from a distribution which is a member of a continuous 1-parameter regular exponential family, so that

$$f(x; \theta) = \exp\{\phi(\theta)h(x) + c(\theta) + d(x)\}, \quad x \in \mathcal{X},$$

and assume that ϕ is an increasing function. Let $T = t(\mathbf{X}) = \sum_{i=1}^n h(X_i)$. Consider $H_0 : \theta = \theta_0$ and for fixed $\alpha > 0$ choose k_α such that $P_{\theta_0}(t(\mathbf{X}) \geq k_\alpha) = \alpha$. The following test is proposed: Reject H_0 if $t(\mathbf{x}) \geq k_\alpha$.

(a) Let $\theta_1 > \theta_0$. Show that the test is most powerful when testing H_0 against $H_1 : \theta = \theta_1$.

(b) Show that the test is uniformly most powerful when testing H_0 against $H_1 : \theta > \theta_0$.

Solution: Under H_0 the likelihood function is

$$L(\theta, \mathbf{x}) = \exp\left\{ \phi(\theta) \sum_{i=1}^{n} h(x_i) + nc(\theta) + \sum_{i=1}^{n} d(x_i) \right\}.$$

(a) For testing H_0 against $H_1 : \theta = \theta_1$ the Neyman-Pearson Lemma tells us that the most powerful test is the likelihood ratio test. The likelihood ratio here is

$$LR = \frac{L(\theta_1, \mathbf{x})}{L(\theta_0, \mathbf{x})} = \exp\left\{ (\phi(\theta_1) - \phi(\theta_0))t(\mathbf{x}) + n(c(\theta_1) - c(\theta_0)) \right\}.$$

Now, for any c,

$$LR \geq c \iff (\phi(\theta_1) - \phi(\theta_0))t(\mathbf{x}) + n(c(\theta_1) - c(\theta_0)) \geq d,$$

where d is some other constant; with more constants f and k we continue to argue:

$$\begin{aligned} & (\phi(\theta_1) - \phi(\theta_0))t(\mathbf{x}) + n(c(\theta_1) - c(\theta_0)) \geq d \\ \iff & \quad (\phi(\theta_1) - \phi(\theta_0))t(\mathbf{x}) \geq f \\ \iff & \quad t(\mathbf{x}) \geq k. \end{aligned}$$

The last step follows because ϕ is assumed to be increasing, and $\theta_1 > \theta_0$. Hence the proposed test is the LR-test, and by the Neyman-Pearson Lemma it is most powerful.

(b) The test does not depend on θ_1, and it it most powerful for all θ_1 which are larger than θ_0, hence it is uniformly most powerful when testing H_0 against $H_1 : \theta > \theta_0$.

4. Consider the linear model

$$Y = \beta x + \gamma z + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, and σ^2 is unknown. Let $\hat{\sigma^2}$ be the maximum-likelihood estimator for σ^2. Let $\theta = (\beta, \gamma)^T$, and denote the estimated variance-covariance matrix of θ by $\hat{V}(\hat{\theta})$; say,

$$\hat{V}(\hat{\theta}) = \hat{\sigma^2} \begin{pmatrix} w_{11} & w_{12} \\ w_{12} & w_{22} \end{pmatrix}.$$

a) Show that for testing $H_0 : \beta\gamma = 1$ the Wald statistic is

$$\frac{(\hat{\beta}\hat{\gamma} - 1)^2}{\hat{\sigma^2}(\hat{\gamma}^2 w_{11} + 2\hat{\beta}\hat{\gamma} w_{12} + \hat{\beta}^2 w_{22})}.$$

b) Derive the Wald statistic for testing $H_0 : \beta = \gamma^{-1}$.

Solution: We use that, to test $H_0 : g(\theta) = 0$, where g is a differentiable function, then the delta method gives the Wald test statistic

$$W = g(\hat{\theta})^T \{ G(\hat{\theta})(I_n(\hat{\theta}))^{-1} G(\hat{\theta})^T \}^{-1} g(\hat{\theta}),$$

where $G(\theta) = \frac{\partial g(\theta)}{\partial \theta}^T$. As $(I_n(\hat{\theta}))^{-1}$ is the approximate variance-covariance matrix for $\hat{\theta}$, we use $\hat{V}(\hat{\theta})$ for $(I_n(\hat{\theta}))^{-1}$.

a) Here, $g(\theta) = \beta\gamma - 1$, so that $G(\theta) = (\gamma, \beta)$, and

$$G(\hat{\theta})(I_n(\hat{\theta}))^{-1}G(\hat{\theta})^T = \hat{\sigma^2}(\hat{\gamma}^2 w_{11} + 2\hat{\gamma}\hat{\beta}w_{12} + \hat{\beta}^2 w_{22}),$$

so that

$$W = (\hat{\beta}\hat{\gamma} - 1)^2 \left(\hat{\sigma^2}(\hat{\gamma}^2 w_{11} + 2\hat{\gamma}\hat{\beta}w_{12} + \hat{\beta}^2 w_{22})\right)^{-1},$$

as required.

b) Here, $g(\theta) = \beta - \gamma^{-1}$, so that $G(\theta) = (1, \gamma^{-2})$, yielding

$$G(\hat{\theta})(I_n(\hat{\theta}))^{-1}G(\hat{\theta})^T = \hat{\sigma^2}(w_{11} + 2\hat{\gamma}^{-2}w_{12} + \hat{\gamma}^{-4}w_{22}),$$

so that

$$W = (\hat{\beta} - \hat{\gamma}^{-1})^2 \left(\hat{\sigma^2}(w_{11} + 2\hat{\gamma}^{-2}w_{12} + \hat{\gamma}^{-4}w_{22})\right)^{-1},$$

which simplifies to

$$W = (\hat{\beta}\hat{\gamma} - 1)^2 \left(\hat{\sigma^2}(\hat{\gamma}^2 w_{11} + 2w_{12} + \hat{\gamma}^{-2}w_{22})\right)^{-1}.$$

We note that the statistics in a) and in b) are not the same.

5. Suppose $Y_i \sim N(\beta_0 + \beta_1 x_i{}^{\psi}, \sigma^2)$ for $i = 1, \ldots, n$, where β_0, β_1, ψ and σ are unknown parameters and where the constants x_i are known. Derive the score test of $H_0 : \psi = 1$ against the alternative $H_1^+ : \psi > 1$.

Solution: The log likelihood is

$$\ell(\psi, \beta, \sigma) = const. - \frac{n}{2} log\sigma^2 - \frac{1}{2\sigma^2}\sum(y_j - \beta_0 - \beta_1 x_j^{\psi})^2$$

and differentiation w.r.t. ψ gives

$$\frac{\partial \ell}{\partial \psi} = \frac{1}{\sigma^2}\beta_1 \sum(y_j - \beta_0 - \beta_1 x_j^{\psi})x_j^{\psi} log x_j.$$

Under H_0 the m.l.e.s $\hat{\beta}_0, \hat{\beta}_1$ are the normal least squares estimates, and

$$\hat{\sigma}^2 = \frac{1}{n}\sum(y_j - \hat{\beta}_0 - \hat{\beta}_1 x_j)^2.$$

The score test statistic is

$$T = \frac{1}{\hat{\sigma}^2}\hat{\beta}_1 \sum(y_j - \hat{\beta}_0 - \hat{\beta}_1 x_j^{\psi})x_j^{\psi} log x_j,$$

and we reject H_0 for large positive values of T.

Let $u_j = x_j log x_j$, then the information matrix works out to be

$$I_n(\psi_0, \beta_0, \beta_1, \sigma) = \sigma^{-2}\begin{pmatrix} \beta_1^2 \sum u_j^2 & \beta_1 \sum u_j & \beta_1 \sum u_j x_j & 0 \\ \beta_1 \sum u_j & n & \sum x_j & 0 \\ \beta_1 \sum u_j x_j & \sum x_j & \sum x_j^2 & 0 \\ 0 & 0 & 0 & 2n \end{pmatrix}$$

and the variance of T is the inverse of the $(1,1)$ element of I_n^{-1}.

You can calculate the inverse; put $s_{xx} = \sum_j (x_j - \overline{x})^2$ and $s_{uu} = \sum_j (u_j - \overline{u})^2$ to obtain that

$$Var(T) \quad = \quad \frac{\sigma^2 s_{xx}}{\beta_1^2 \left\{ s_{xx}s_{uu} - (\sum u_j x_j - n\overline{x}\overline{u})^2 \right\}}.$$

6. Suppose $X_i \sim N(\mu_x, \sigma_x^2)$ for $i = 1, \ldots, n$, and $Y_j \sim N(\mu_y, \sigma_y^2)$ for $j = 1, \ldots, m$, where μ_x, μ_y, σ_x and σ_y are all unknown. Let S_x^2 and S_y^2 denote the sample variances. Use the pivot $(S_x^2/\sigma_x^2)/(S_y^2/\sigma_y^2)$ to obtain an exact upper $1-\alpha$ confidence limit for $\psi = \sigma_y^2/\sigma_x^2$. How would you construct the corresponding confidence limit for ψ if both μ_x and μ_y were known?

Solution: We know that

$$\frac{S_x^2 \sigma_y^2}{S_y^2 \sigma_X^2} \sim \frac{\chi^2_{n-1}/(n-1)}{\chi^2_{m-1}/(m-1)} = F_{n-1,m-1}$$

(see the book by Rice) and hence an exact upper $1-\alpha$ confidence bound for $\psi = \frac{\sigma_Y^2}{\sigma_X^2}$ is

$$\frac{s_y^2}{s_x^2} F_{n-1,m-1}(1-\alpha).$$

If both μ_x and μ_y are known then we use

$$\hat{\sigma}_X^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu_x)^2, \quad \hat{\sigma}_Y^2 = \frac{1}{m}\sum_{i=1}^m (Y_i - \mu_Y)^2,$$

and $F_{n,m}$.

7. Let $X_{(n)}$ be the largest value in a sample of size n drawn from the uniform distribution on $[0, \theta]$. Show that $X_{(n)}/\theta$ is a pivot. Using this pivot, find a $100(1-\alpha)\%$ confidence interval for θ. Discuss how you would test the hypothesis that θ takes a specific value θ_0 for such a sample.

Solution: We know that

$$P(X_{(n)} \le x) = (P(X_i \le x))^n = \theta^{-n} x^n, \quad 0 \le x \le \theta,,$$

and therefore

$$P(X_{(n)}/\theta \le x) = x^n, \quad 0 \le x \le 1,$$

and so $X_{(n)}/\theta$ is a pivot. For a $100(1-\alpha)\%$ confidence interval we would like

$$1-\alpha = P\left(a \le X_{(n)}/\theta < b\right) = P\left(\frac{X_{(n)}}{b} < \theta < \frac{X_{(n)}}{a}\right).$$

If we choose $b = 1$ then we find $1 - a^n = 1 - \alpha$, yielding $a = \alpha^{1/n}$. Note that other choices of b are possible. For a hypothesis test, we would accept $H_0 : \theta = \theta_0$ if $\theta_0 < X_{(n)}$, or if $\theta_0 > X_{(n)}\alpha^{-1/n}$.