NTHU MATH 2820 Final Exam Solution Jun 17, 2008

1. (*10pts, 2pts for each*)

   (a) False. The probability that $(-2.0, 3.0)$ contains μ is either 0 or 1. The 99% means that if the procedure is repeated over and over to generate many different confidence intervals, about 99% of them will cover μ.

   (b) False. The null would be rejected because $-3 \notin (-2.0, 3.0)$.

   (c) False. Under the (Frequentist) theoretical framework of hypothesis testing, the null hypothesis is either true or false. It is meaningless to talk about "the probability that null hypothesis is ture".

   (d) True.

   (e) True

2. (*15pts, 3pts for each*)

   (a) **(1)** $X \sim \text{Binomial}(1919, p)$ (Note that the total number of deaths, i.e., 1919, is treated as a fixed number, rather than a random variable.)

   (2) $\Omega = \{p|\ 0 < p \le 1/2\}$ (**Note.** It is either there is no holiday effect, i.e., $p = 1/2$, or there is a holiday effect, i.e., $0 < p < 1/2$. We do not assume there exists an "anti-holiday" effect, i.e., more people dying before holiday than after holiday.)

   (3) $\Omega_0 = \{1/2\}$

   (b) **(1)** $X_1, \ldots, X_{20}$ are i.i.d. from Exponential(λ)

   (2) $\Omega = \{\lambda|\ 0 < \lambda < \infty\}$

   (3) From the customers viewpoint, we would like to protect the hypothesis "the mean time to failure is less than 2 years" and therefore use it as the null hypothesis, i.e., $\Omega_0 = \{\lambda|\ \lambda \ge \frac{1}{2\times 365} = \frac{1}{730}\}$. (**Note**. $E(X_i) = 1/\lambda$.)

   (c) **(1)** $X_i \sim \text{Binomial}(200, p_i)$, $i = 1, \ldots, 4$, and $X_1, \ldots, X_4$ are independent

   (2) $\Omega = \{(p_1, \ldots, p_4)|\ 0 < p_i < 1,\ i = 1, \ldots, 4\}$. (Note that $\dim(\Omega) = 4$, not 3.)

   (3) $\Omega_0 = \{(p_1, \ldots, p_4)|\ 0 < p_1 = \cdots = p_4 < 1\}$. (Note that $\dim(\Omega_0) = 1$.)

   (d) **(1)** $X_1, \ldots, X_8$ are i.i.d. from Normal(μ, σ^2)

   (2) $\Omega = \{(\mu, \sigma)| -\infty < \mu < \infty,\ 0 < \sigma < \infty\}$

   (3) $\Omega_0 = \{(\mu, \sigma)|\ \mu = 0,\ 0 < \sigma < \infty\}$

   (e) **(1)** $(X_0, X_1, \ldots, X_5) \sim \text{Multinomial}(280, p_0, p_1, \ldots, p_5)$

   (2) $\Omega = \{(p_0, p_1, \ldots, p_5)|\ \sum_{i=0}^{5} p_i = 1\}$ (Note that $\dim(\Omega) = 6 - 1 = 5$.)

   (3) $\Omega_0 = \{(p_0, p_1, \ldots, p_5)|\ p_i = P(Z = i),\ i = 0, 1, \ldots, 5$, where $Z \sim \text{Binomial}(5, p)$ and $0 < p < 1\ \}$ (Note that $\dim(\Omega_0)$=1.)

3. (a) (*15pts*) The joint pmf of $X_1, \ldots, X_n$ is

$$f(\mathbf{x}, \lambda) = e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i} \cdot \left(\prod_{i=1}^n \frac{1}{x_i!} \right).$$

Therefore, the likelihood ratio is

$$\Lambda = \frac{f(\mathbf{x}, \lambda_0)}{f(\mathbf{x}, \lambda_1)} = e^{n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1} \right)^{\sum_{i=1}^n x_i}.$$

Because $\lambda_1 > \lambda_0$, Λ decreases as $\sum_{i=1}^n X_i$ increases. A randomized test function based on the likelihood ratio is then given by:

$$\phi(\mathbf{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i > c \\ \gamma, & \text{if } \sum_{i=1}^n X_i = c \\ 0, & \text{if } \sum_{i=1}^n X_i < c \end{cases}.$$

The c and γ are determined by

$$\mathrm{E}_{\lambda_0}(\phi) = P\left(\sum_{i=1}^n X_i > c \right) + \gamma \cdot P\left(\sum_{i=1}^n X_i = c \right) = \alpha,$$

where $\sum_{i=1}^n X_i \sim P(n\lambda_0)$.

(b) (*5pts*) By Neyman-Pearson lemma, the test in (a) is the most powerful test for any particular simple alternative $H_A : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$. Furthermore, because the rejection region, i.e., c and γ, of the test does not depend on λ_1, the test is UMP for $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda > \lambda_0$.

4. (a) (*7pts*) Because $\Omega_0 = \{\theta_0\}$ and $\Omega = (0, \infty)$,

$$\Lambda(x_1, \ldots, x_n) = \frac{\max_{\theta \in \Omega_0} \mathcal{L}(\theta, \mathbf{x})}{\max_{\theta \in \Omega} \mathcal{L}(\theta, \mathbf{x})} = \frac{\left[\frac{1}{\theta_0} \right]^n \cdot I_{[0,\theta_0]}(x_{(n)})}{\left[\frac{1}{x_{(n)}} \right]^n \cdot I_{[0,x_{(n)}]}(x_{(n)})} = \left[\frac{x_{(n)}}{\theta_0} \right]^n \cdot I_{[0,\theta_0]}(x_{(n)}).$$

(b) (*4pts*) As a function of $x_{(n)}$, Λ increases from 0 to 1 when $x_{(n)}$ increases from 0 to θ_0, and $\Lambda = 0$ when $x_{(n)}$ is larger than θ_0. Therefore, for $0 < s < 1$,

$$\begin{aligned} \Lambda(X_1, \ldots, X_n) < s \quad &\Leftrightarrow \quad X_{(n)} > \theta_0 \text{ or } \left[\frac{X_{(n)}}{\theta_0} \right]^n < s \\ &\Leftrightarrow \quad X_{(n)} > \theta_0 \text{ or } X_{(n)} < s^{1/n}\theta_0. \qquad \text{(I)} \end{aligned}$$

(c) (*9pts*) The rejection region of the GLR test is $\Lambda < s$, where s is determined by

$$\begin{aligned} \alpha &= P(\Lambda < s | H_0) = P\left(\{X_{(n)} > \theta_0\} \cup \{X_{(n)} < s^{1/n}\theta_0\} \middle| H_0 \right) \\ &= P\left(X_{(n)} > \theta_0 \middle| \theta = \theta_0 \right) + P\left(X_{(n)} < s^{1/n}\theta_0 \middle| \theta = \theta_0 \right) \\ &= \int_{\theta_0}^{\infty} 0 \, dx + \int_0^{s^{1/n}\theta_0} \frac{nx^{n-1}}{\theta_0^n} \, dx = 0 + \left. \frac{x^n}{\theta_0^n} \right|_0^{s^{1/n}\theta_0} = s. \end{aligned}$$

Therefore, for $\alpha = 0.05$, we can substitute $s = 0.05$ into (I) to get the rejection region $\left\{X_{(n)} < (0.05)^{1/n}\theta_0\right\} \cup \left\{X_{(n)} > \theta_0\right\}$.

(d) (*4pts*) The acceptance region is $(0.05)^{1/n}\theta_0 \le X_{(n)} \le \theta_0$, i.e.,

$$\begin{aligned} 0.95 &= P\left((0.05)^{1/n}\theta_0 \le X_{(n)} \le \theta_0 \middle| H_0\right) \\ &= P\left(\frac{(0.05)^{1/n}}{X_{(n)}} \le \frac{1}{\theta_0} \le \frac{1}{X_{(n)}} \middle| \theta = \theta_0\right) \\ &= P\left(X_{(n)} \le \theta_0 \le \frac{X_{(n)}}{(0.05)^{1/n}} \middle| \theta = \theta_0\right) \end{aligned}$$

Therefore, $\left[X_{(n)}, 20^{1/n}X_{(n)}\right]$ is a 95% confidence interval for θ.

5. (a) (*2pts*) $\dim(\Omega_0) = 1$ and $\dim(\Omega) = m$

(b) (*8pts*) Because

$$\begin{aligned} \Lambda(x_1, \ldots, x_m) &= \frac{\max_{\theta \in \Omega_0} \mathcal{L}(\theta, \mathbf{x})}{\max_{\theta \in \Omega} \mathcal{L}(\theta, \mathbf{x})} = \frac{\prod_{i=1}^{m} \binom{n_i}{x_i} \hat{p}^{x_i}(1-\hat{p})^{n_i - x_i}}{\prod_{i=1}^{m} \binom{n_i}{x_i} \hat{p}_i^{x_i}(1-\hat{p}_i)^{n_i - x_i}} \\ &= \prod_{i=1}^{m} \left(\frac{\hat{p}}{\hat{p}_i}\right)^{x_i} \left(\frac{1-\hat{p}}{1-\hat{p}_i}\right)^{n_i - x_i}, \end{aligned}$$

we can get

$$\begin{aligned} -2\log\Lambda &= -2\sum_{i=1}^{m}\left[x_i \log\left(\frac{\hat{p}}{\hat{p}_i}\right) + (n_i - x_i)\log\left(\frac{1-\hat{p}}{1-\hat{p}_i}\right)\right] \\ &= -2\sum_{i=1}^{m}\left[n_i\hat{p}_i \log\left(\frac{n_i\hat{p}}{n_i\hat{p}_i}\right) + n_i(1-\hat{p}_i)\log\left(\frac{n_i(1-\hat{p})}{n_i(1-\hat{p}_i)}\right)\right] \\ &= -2\sum_{i=1}^{m}\sum_{j=1}^{2} O_{ij}\log\left(\frac{E_{ij}}{O_{ij}}\right) = 2\sum_{i=1}^{m}\sum_{j=1}^{2} O_{ij}\log\left(\frac{O_{ij}}{E_{ij}}\right). \end{aligned}$$

(c) (*2pts*) Because $\dim(\Omega) - \dim(\Omega_0) = m - 1$, the large sample distribution of $-2\log\Lambda$ is Chi-square distribution with degrees of freedom $m-1$, i.e., χ^2_{m-1}.

(d) (*4pts*) The rejection region is $2\sum_{i=1}^{m}\sum_{j=1}^{2} O_{ij}\log\left(\frac{O_{ij}}{E_{ij}}\right) > c$, where c is determined by $P(\chi^2_{m-1} > c) = \alpha$.

6. (a) (*5pts*) The posterior pdf is

$$h(\theta | x_1, \ldots, x_n) \propto f(x_1, \ldots, x_n | \theta) \cdot g(\theta) = \left[\prod_{i=1}^{n} f(x_i | \theta)\right] \cdot g(\theta)$$

$$
\begin{aligned}
&\propto \left[\prod_{i=1}^{n} \theta^{x_i} e^{-\theta}\right] \cdot \theta^{\alpha-1} e^{-\lambda\theta} \\
&= \theta^{\sum_{i=1}^{n} x_i} e^{-n\theta} \cdot \theta^{\alpha-1} e^{-\lambda\theta} \\
&= \theta^{(n\overline{X}+\alpha)-1} e^{-(n+\lambda)\theta},
\end{aligned}
$$

which follows the form of the pdf of Gamma distribution with shape parameter $n\overline{X}+\alpha$ and scale parameter $n+\lambda$, i.e., $\Theta|x_1,\ldots,x_n \sim \Gamma(n\overline{X}+\alpha, n+\lambda)$.

(b) (*2pts*) Yes, because both the prior and the posterior distributions belong to Gamma family when the sample is from Poisson distribution.

(c) (*5pts*) Because $\Theta|x_1,\ldots,x_n \sim \Gamma(n\overline{X}+\alpha, n+\lambda)$, the Bayes estimator is

$$
\mu_{\text{post}} = E[\Theta|x_1,\ldots,x_n] = \frac{n\overline{X}+\alpha}{n+\lambda} = \frac{n}{n+\lambda}\cdot\overline{X} + \frac{\lambda}{n+\lambda}\cdot\frac{\alpha}{\lambda}, \qquad \text{(II)}
$$

where $\frac{\alpha}{\lambda}$ is the prior mean, and sum of the weights is one, i.e., $\frac{n}{n+\lambda}+\frac{\lambda}{n+\lambda}=1$.

(d) (*3pts*) When n is large, the weights in (II) will approximate 1 and 0 respectively, i.e.,

$$
\frac{n}{n+\lambda} \approx 1 \quad \text{and} \quad \frac{\lambda}{n+\lambda} \approx 0.
$$

Therefore, $\mu_{\text{post}} \approx \overline{X}$, which is a function of sample.