UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 9- Solutions
Fall 2014

Issued: Oct 31
Due: Thursday, Nov 6

Problem 9.1

(a) The likelihood ratio test statistic λ is equal to:

$$\lambda(X) = \frac{\sup_{\theta\in\Omega_1} p_\theta(X)}{\sup_{\theta\in\Omega_0} p_\theta(X)} = \begin{cases} \exp\{-\frac{1}{2}\inf_{\theta\in\Omega_1}\|X-\theta\|_2^2\} & X\in\Omega_0 \\ \exp\{\frac{1}{2}\inf_{\theta\in\Omega_0}\|X-\theta\|_2^2\} & X\in\Omega_1 \end{cases}$$

Note that for the first case, $\lambda \leq 1$, we will not reject the hypothesis and meanwhile, $D = 0$. In the second case, rejecting when $\lambda > k$ is equivalent to reject when $D = 2\log\lambda > \sqrt{2\log k}$. Therefore, a test with $I\{\lambda > k\}$, $k > 1$ is equivalent to $I\{D > \sqrt{2\log k}\}$.

(b) For any $\theta \in \Omega_0$, we have

$$E_\theta I\{D > c\} = \int_S \frac{1}{2\pi}\exp\{-\frac{1}{2}\left((x_1-\theta_1)^2 + (x_2-\theta_2)^2\right)\}d\theta$$

where $S = \inf_{\theta\in\Omega_0}\|x-\theta\| > c$ does not depend on θ and $S\cap\Omega_0 = \emptyset$. The supremum is obtained at $\theta = 0$ then

$$\begin{aligned}\alpha &= \int_{\{x_1<0,x_2>c\}\cup\{x_1>c,x_2<0\}\cup\{x_1^2+x_2^2>c^2,x_1,x_2>0\}} \frac{1}{2\pi}\exp\{-\frac{1}{2}\left(x_1^2+x_2^2\right)\}d\theta \\ &= 1-\Phi(c) + \frac{1}{4}e^{-c^2/2}.\end{aligned}$$

Problem 9.2

(a) Under H_0, $\bar{Y}_n \sim \mathcal{N}(\mu_0, \sigma^2/n)$ thus $t_\alpha = \mu_0 + \sigma\Phi^{-1}(1-\alpha)/\sqrt{n}$.

(b) The power of the test is

$$\begin{aligned}
\beta(\mu_1) &= P_{\mu_1}(\bar{Y}_n \geq t_\alpha) \\
&= P_{\mu_1}(\sqrt{n}/\sigma(\bar{Y}_n - \mu_1) \leq \sqrt{n}/\sigma(t_\alpha - \mu_1)) \\
&= P_{\mu_1}(\sqrt{n}/\sigma(\bar{Y}_n - \mu_1) \leq \sqrt{n}/\sigma(\mu_1 - t_\alpha)) \\
&= \Phi(\sqrt{n}/\sigma(\mu_1 - t_\alpha)) \\
&= \Phi(z_\alpha + \delta_\alpha).
\end{aligned}$$

The power of test increases/decreases as $\mu_1 - \mu_0$ increases/decreases, intuitively, it becomes easier to differentiate H_0 from H_1 as $\mu_0 - \mu_1$ increases. The power of test increases/decreases as σ decreases/increases, intuitively, it becomes easier to differentiate H_0 from H_1 as noise level goes down. The power of test increases as n increases, intuitively, we are getting more evidence.

(c) Under H_1, $P_{\mu_1}(Z_i = 1) = P_{\mu_1}(Y_i > \mu_0) = \Phi(\delta_n/\sqrt{n})$. Therefore, $n\bar{Z}_n \sim Bin(n, \Phi(\delta_n/\sqrt{n}))$. By CLT,

$$\sqrt{n}(\bar{Z}_n - p_{\mu_1}) \xrightarrow{d} \mathcal{N}(0, p_{\mu_1}(1 - p_{\mu_1}))$$

where $p_{\mu_1} := \Phi(\delta_n/\sqrt{n})$.

Meanwhile, under H_0,

$$\sqrt{n}(\bar{Z}_n - \frac{1}{2}) \xrightarrow{d} \mathcal{N}(0, \frac{1}{4})$$

hence

$$P_{\mu_0}(\bar{Z}_n \geq \frac{1}{2} - \frac{z_\alpha}{2\sqrt{n}}) \to \alpha.$$

The power function

$$\beta(\mu_1) = P_{\mu_1}(\bar{Z}_n \geq \frac{1}{2} - \frac{z_\alpha}{2\sqrt{n}}) \to \Phi(\frac{np_{\mu_1} - n/2 + \sqrt{n}z_\alpha/2}{\sqrt{np_{\mu_1}(1 - p_{\mu_1})}})$$

where $p_{\mu_1} := \Phi(\delta_n/\sqrt{n})$.

(d) Plug in the approximation $\Phi(\delta_n/\sqrt{n})$ by $1/2 + \frac{\delta_n}{\sqrt{2\pi n}}$ then the power function

$$\beta(\mu_1) \approx \Phi(\frac{\delta_n/\sqrt{2\pi} + z_\alpha/2}{\sqrt{1/4 - \delta_n^2/2\pi n}}) \to \Phi(z_\alpha + \sqrt{2/\pi}\delta_n).$$

Problem 9.3

(a) Under H_0, $I(Y_i > \mu_0) \sim Bin(1, 1/2)$, thus $S \sim Bin(n, 1/2)$. Based on the distribution of S, we could find k_n and test $\phi(S) = I(S > k_n)$ st:

$$P(S > k_n) \leq \alpha.$$

Note that $\phi(S)$ does not depend on the exact form of F.

(b) Approximate binomial distribution by normal distribution,

$$\alpha(S) = P_0(S > s) = P_0((S - n/2)/\sqrt{n/4} > (s - n/2)/\sqrt{n/4}) = \Phi(\sqrt{n} - 2s/\sqrt{n}).$$

Problem 9.4

a) The action space is given by $\mathcal{A} = \{0, 1\}$. For $\theta = \theta_0$, the loss function is given by:

$$l(\theta_0, \delta) = \begin{cases} 0, & \text{if } \delta = 0 \\ 1, & \text{if } \delta = 1 \end{cases}$$

For $\theta = \theta_1$, the loss function is given by:

$$l(\theta_1, \delta) = \begin{cases} 0, & \text{if } \delta = 1 \\ 1, & \text{if } \delta = 0 \end{cases}$$

It follows that $\mathbb{E}\left(l(\theta, \delta(X)) | \theta = \theta_0\right) = \mathbb{E}\left(\delta(X) | \theta = \theta_0\right)$ and $\mathbb{E}\left(l(\theta, \delta(X)) | \theta = \theta_1\right) = \mathbb{E}\left(1 - \delta(X) | \theta = \theta_1\right)$ As a result:

$$\begin{aligned} r(\lambda, \delta) &= \mathbb{E}\left(l(\theta, \delta(X))\right) = \mathbb{E}\left(l(\theta, \delta(X)) | \theta = \theta_0\right) \mathbb{P}(\theta = \theta_0) + \mathbb{E}\left(l(\theta, \delta(X)) | \theta = \theta_1\right) \mathbb{P}(\theta = \theta_1) \\ &= \lambda_0 \mathbb{E}_0\left(\delta(X)\right) + (1 - \lambda_0) \mathbb{E}_1\left(1 - \delta(X)\right) \end{aligned}$$

b) To minimize the Bayes risk, it is sufficient to minimize the posterior risk. The posterior risk of taking action δ is given by:

$$\begin{aligned} \mathbb{E}\left(L(\theta, \delta) \middle| X\right) &= \delta \frac{\lambda_0 \mathbb{P}(X|\theta_0)}{\lambda_0 \mathbb{P}(X|\theta_0) + (1 - \lambda_0) \mathbb{P}(X|\theta_1)} + (1 - \delta) \frac{(1 - \lambda_0) \mathbb{P}(X|\theta_1)}{\lambda_0 \mathbb{P}(X|\theta_0) + (1 - \lambda_0) \mathbb{P}(X|\theta_1)} \\ &= \delta \left[\frac{\lambda_0 \mathbb{P}(X|\theta_0) - (1 - \lambda_0) \mathbb{P}(X|\theta_1)}{\lambda_0 \mathbb{P}(X|\theta_0) + (1 - \lambda_0) \mathbb{P}(X|\theta_1)} \right] + \frac{(1 - \lambda_0) \mathbb{P}(X|\theta_1)}{\lambda_0 \mathbb{P}(X|\theta_0) + (1 - \lambda_0) \mathbb{P}(X|\theta_1)} \end{aligned}$$

Hence, the optimal decision $\delta^*(X)$ is given by:

$$\begin{aligned} \delta^*(X) &= \mathbb{I}\left((1 - \lambda_0) \mathbb{P}(X|\theta_1) - \lambda_0 \mathbb{P}(X|\theta_0) \geq 0\right) \\ &= \mathbb{I}\left(\frac{\mathbb{P}(X|\theta_1)}{\mathbb{P}(X|\theta_0)} \geq \frac{\lambda_0}{1 - \lambda_0}\right) \end{aligned}$$

c) From item b:

$$\delta^*(X) \quad = \quad \mathbb{I}\left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mathbb{P}(X_i|\theta_1)}{\mathbb{P}(X_i|\theta_0)}\right) \geq \frac{1}{n}\log\left(\frac{\lambda_0}{1-\lambda_0}\right)\right]$$

Now, under H_0:

$$\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mathbb{P}(X_i|\theta_1)}{\mathbb{P}(X_i|\theta_0)}\right) \xrightarrow{p} \mathbb{E}_{\theta_0}\left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mathbb{P}(X_i|\theta_1)}{\mathbb{P}(X_i|\theta_0)}\right)\right] = D(\theta_0||\theta_1) > 0$$

while, under H_1:

$$\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mathbb{P}(X_i|\theta_1)}{\mathbb{P}(X_i|\theta_0)}\right) \xrightarrow{p} \mathbb{E}_{\theta_1}\left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mathbb{P}(X_i|\theta_1)}{\mathbb{P}(X_i|\theta_0)}\right)\right] = -D(\theta_1||\theta_0) < 0$$

Simultaneously, we have that for any $\lambda_0 \in (0,1)$:

$$\frac{1}{n}\log\left(\frac{\lambda_0}{1-\lambda_0}\right) \quad \to \quad 0$$

Hence:

$$\delta^*(X) \xrightarrow{p} \begin{cases} 1, & \text{under } H_1 \\ 0, & \text{under } H_0 \end{cases}$$

Now, from boundedness of δ and the convergence in probability above, we get:

$$\mathbb{E}_0[\delta(X)] \quad \to \quad 0$$

and

$$\mathbb{E}_1[1-\delta(X)] \quad \to \quad 0$$

which in turn yield:

$$r(\lambda, \delta_\lambda) \quad = \quad \lambda\mathbb{E}_0[\delta(X)] + (1-\lambda)(1-\mathbb{E}_1[\delta(X)]) \to 0$$

as required.

Problem 9.5

(a) For $\theta = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2)$,

$$\sup_{\theta\in\Omega} l(\theta) \quad = \quad \frac{n}{2}\log\hat{\sigma}_x^2 - \frac{n}{2}\log\hat{\sigma}_y^2 - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(X_i-\hat{\mu}_x)^2 - \frac{1}{2\hat{\sigma}_y^2}\sum_{i=1}^{n}(Y_i-\hat{\mu}_y)^2 - n\log(2\pi)$$

$$\sup_{\theta\in\Omega_0} l(\theta) \quad = \quad -n\log\hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2}\left(\sum_{i=1}^{n}(X_i-\hat{\mu})^2 + \sum_{i=1}^{n}(Y_i-\hat{\mu})^2\right) - n\log(2\pi)$$

where $\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^n X_i$, $\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^n Y_i$, $\hat{\mu} = \frac{1}{2n}\sum_{i=1}^n (X_i + Y_i)$, $\hat{\sigma}_x^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \hat{\mu}_x)^2$, $\hat{\sigma}_y^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \hat{\mu}_y)^2$, $\hat{\sigma}^2 = \frac{1}{2n}\left(\sum_{i=1}^n (X_i - \hat{\mu})^2 + \sum_{i=1}^n (Y_i - \hat{\mu})^2\right)$ Thus,

$$2\log\left(\sup_{\theta\in\Omega} l(\theta)\right) - 2\log\left(\sup_{\theta\in\Omega_0} l(\theta)\right) \xrightarrow{d} \chi_2^2$$

(b)

$$\begin{aligned}
\sup_{\theta\in\Omega} l(\theta) &= \sum_{i=1}^k \left(-n\hat{\theta}_i + \log\hat{\theta}_i \sum_{j=1}^n X_{ij}\right) \\
\sup_{\theta\in\Omega_0} l(\theta) &= -kn\hat{\theta} + \log\hat{\theta}\sum_{i=1}^k\sum_{j=1}^n X_{ij}
\end{aligned}$$

where $\hat{\theta}_i = \frac{1}{n}\sum_{j=1}^n X_{ij}$, $i = 1, \dots, k$, $\hat{\theta} = \frac{1}{kn}\sum_{i=1}^k\sum_{j=1}^n X_{ij}$ Thus,

$$2\log\left(\sup_{\theta\in\Omega} l(\theta)\right) - 2\log\left(\sup_{\theta\in\Omega_0} l(\theta)\right) \xrightarrow{d} \chi_{k-1}^2$$

(c)

$$\begin{aligned}
\sup_{\theta\in\Omega} l(\theta) &= n\log\hat{\theta} - \hat{\theta}\sum_{i=1}^n X_i + n\log\hat{\mu} - \hat{\mu}\sum_{i=1}^n Y_i \\
\sup_{\theta\in\Omega_0} l(\theta) &= n\log 2 + 2n\log\hat{\theta}_0 - \hat{\theta}_0\left(\sum_{i=1}^n X_i + 2\sum_{i=1}^n Y_i\right)
\end{aligned}$$

where $\hat{\theta} = \frac{n}{\sum_{i=1}^n X_i}$, $\hat{\mu} = \frac{n}{\sum_{i=1}^n Y_i}$, $\hat{\theta}_0 = \frac{2n}{\sum_{i=1}^n X_i + 2\sum_{i=1}^n Y_i}$ Thus,

$$2\log\left(\sup_{\theta\in\Omega} l(\theta)\right) - 2\log\left(\sup_{\theta\in\Omega_0} l(\theta)\right) \xrightarrow{d} \chi_1^2$$