Statistical Inference 553. Spring 2003

Homework 5

April, 2003

1. **Exercise 8.10:** Given a gamma(α,β) prior, from exercise 7.24, we know that the posterior distribution follows a gamma(α',β') where

$$\alpha' = \alpha + \sum_{i=1}^{n} X_i; \beta' = \beta/(n\beta + 1)$$

   (a) Expressions for the posterior probabilities of H_0 and H_1 are:

$$P(H_0|\underline{x}) = \int_0^{\lambda_0} Gamma(\alpha', \beta')d\lambda; P(H_1|\underline{x}) = \int_{\lambda_0}^{\infty} Gamma(\alpha', \beta')d\lambda$$

   where $Gamma(\alpha', \beta')$ represents a gamma pdf with parameters α' and β'.

   (b) With this prior, the parameters for the posterior distribution for λ are:

$$\alpha' = (5 + 2\sum_{i=1}^{n} X_i)/2; \beta' = 2/(2n + 1)$$

   Define $Y = (2n + 1)\lambda$. Via "change of variable", Y has a gamma pdf with parameters α' and $\beta = 2$, i.e. a chi-square distribution with $p = 5 + 2\sum_{i=1}^{n} X_i$ degrees of freedom. Then,

$$P(H_0|\underline{x}) = P(\lambda \le \lambda_0|\underline{x}) = P(Y \le (2n + 1)\lambda_0)$$

   Since Y is a chi-square random variable, we only need find the cumulative probability up to the value $(2n+1)\lambda_0$ under the chi-square distribution table with p degrees of freedom.

2. **Exercise 8.14:** We reject the null hypothesis if $\sum_{i=1}^{n} X_i > c$. Under H_0, we wish to have:

$$Pr[\sum_{i=1}^{n} X_i > c|\theta = 0.49] = 0.01$$

   With the CLT approximation, $Z = (\sum X_i - n(0.49))/\sqrt{n(0.49)(0.51)}$ is approximately a $N(0, 1)$. Then,

$$c = n(0.49) + (2.33)(0.4999)\sqrt{n} \quad (1)$$

   Under H_1, we wish to have:$Pr[\sum_{i=1}^{n} X_i > c|\theta = 0.51] = 0.99$. Then,

$$c = n(0.51) - (2.33)(0.4999)\sqrt{n} \quad (2)$$

If we find n that satisfies (1) and (2), we obtain that:

$$n = 13566.82$$

If we want to detect such a small difference with so small errors, we will need a pretty large sample size!

3. **Exercise 8.16:**

   (a) Since we always reject H_0, $Pr[Reject\ H_0|\theta] = 1$ for any value of θ. Then, the type I error is always 1 for $\theta \in \Theta_0$ and the type II is always 0 for $\theta \in \Theta_1$.

   (b) Since we always accept H_0, $Pr[Reject\ H_0|\theta] = 0$ for any value of θ. Then, the type I error is always 0 and the type II is always 1.

4. **Exercise 8.18:**

   (a) The power is the probability of rejecting H_0 given any value of the parameter. Given the rejection region for this problem, the power function is:

   $$\beta(\theta) = Pr[\bar{X} > \theta_0 + c(\sigma/\sqrt{n})\ or\ \bar{X} < \theta_0 - c(\sigma/\sqrt{n})|\theta]$$

   This is equivalent to:

   $$\beta(\theta) = Pr[Z > c + \sqrt{n}(\theta_0 - \theta)/\sigma\ or\ Z < -c + \sqrt{n}(\theta_0 - \theta)/\sigma]$$

   where $Z = \frac{\bar{X}-\theta}{\sigma/\sqrt{n}}$ follows a $N(0,1)$ distribution. Then,

   $$\beta(\theta) = 1 - \Phi\left(c + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right) + \Phi\left(-c + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right)$$

   where $\Phi(\cdot)$ is the CDF of a $N(0,1)$.

   (b) If a type I error probability of 0.05 is desired, c is the 0.975 quantile of N(0,1) distribution, i.e. $c = 1.96$. If we want a type II error probability of at most 0.25 at $\theta_0 + \sigma$, this implies that

   $$\Phi(1.96 - \sqrt{n}) - \Phi(-1.96 - \sqrt{n}) \leq 0.25$$

   Notice that for $n \geq 1$, $\Phi(-1.96 - \sqrt{n}) \approx 0$. All we need is to choose n such that

   $$\Phi(1.96 - \sqrt{n}) \leq 0.25$$

   .

   This implies that $1.96 - \sqrt{n} \leq -0.67$ or $n \geq 6.92$. With $n = 7$ and $c = 1.96$, we have the required conditions.

5. **Exercise 8.21:** All there is to this exercise is to re-write the proof in page 389. Particularly, you need to change the integral sign for a summation sign on equation (8.3.3). The rest of the details are also applicable to the discrete random variable case.

6. **Exercise 8.22:**

(a) Based on the Neyman-Pearson Lemma, you may show that the UMP is given by the condition: Reject H_0 if and only if $\sum_{i=1}^{n} X_i \leq c$. Under H_0, $\sum_{i=1}^{n} X_i$ follows a Binomial $(n, 0.5)$. Since $\alpha = 0.0547$, using the Binomial CDF, $c = 2$. To obtain the power of this test, we need to compute the following Binomial probability with $n = 10$ and $p = 0.25$.

$$Pr[\sum_{i=1}^{n} X_i \leq 2|p = 0.25] = 0.5256$$

(b) For different values of p, we need to compute:

$$\beta(p) = P[\sum_{i=1}^{10} X_i \geq 6|p]$$

for a Binomial$(10, p)$
Here are some values that you may use to sketch this function: $\beta(0.1) = 0.0001$, $\beta(0.2) = 0.0064$, $\beta(0.25) = 0.197$, $\beta(0.5) = 0.377$, $\beta(0.6) = 0.6331$, $\beta(0.8) = 0.9672$.
Certainly, the size of the test is $\alpha = \beta(0.5) = 0.377$.

(c) It will only be for the cumulative probabilities obtained for a Binomial$(10, 0.5)$. These probabilities are: 0.0010, 0.0107, 0.0547, 0.1719, 0.3770, 0.6230, 0.8281, 0.9453, 0.9893, 0.9990, 1.0.