STA4322 Exam3 (closed book, 9:30 - 10:45, 8/6/2009) Name________________________

1. In hypothesis testing, we need two hypotheses H_0 and H_a, a rejection region RR, and two types of errors α and β. State the two main relations between α and β and the other terms. (10%)

   Sol: $\alpha = P\{RR|H_0\}$ and $1 - \beta = P\{RR|H_a\}$

2. (i) When we say that $\hat{\theta}$ is a MVUE (minimum variance unbiased estimator) for an unknown parameter θ, what do we means. Explain it using commonly used statistical notation. (10%)

   Sol: (i) $E[\hat{\theta}] = \theta$ and for and other unbiased estimator $\tilde{\theta}$, $Var(\tilde{\theta}) \geq Var(\hat{\theta})$

   (ii) What we say that a test is the uniformly most powerful test in hypothesis testing, what do we mean? Use an example to explain this. (10%)

   Sol: Let $\mathrm{H}_0 : p = p_0$ and $\mathrm{H}_a : p > p_0$. Then the uniformly most powerful test is the most powerful for test for all the simple hypothesis $\mathrm{H}_a : p = p_a$ with $p_a > p_0$.

3. The minimum requirement for the hardness index of an alloy is 65. Fifty specimen of the alloy was collected and the sample mean is 68 with sample standard deviation 8. Can the company claim that their alloy meets the hardness requirement? Make you conclusion by p-value. (15%)

   Sol: To guarantee the product meets the requirement, we need to put H_a as $\mu > 65$. Thus H_0 is $\mu \leq 65$.

$$\begin{aligned} p = value &= P(\bar{y} \geq 68) \\ &= P(Z \geq \frac{68 - 65}{8/\sqrt{50}}) \\ &= P(Z \geq 2.65) = 0.004 \end{aligned}$$

4. Let y_1, y_2,..., y_n be a random sample from a Poisson distribution with unknown mean λ. Find the most powerful test for $H_0 : \lambda = 1$ against $H_a : \lambda = 2$. You are allowed to have one threshold to be determined by the type I error probability. (10%)

Sol:

$$l(y_1, y_2, ..., y_n) = \frac{\lambda^{y_1+y_2+...+y_n} e^{-n\lambda}}{y_1! y_2! ... y_n!}$$

Thus, most the powerful test is

$$\frac{l(y_1, y_2, ..., y_n|\lambda = 1)}{l(y_1, y_2, ..., y_n|\lambda = 2)} < k$$

$$\Rightarrow \; 2^{-\sum y_i} e^{-n} < k$$

$$\Rightarrow \; \sum y_i > k'$$

5. A study was conducted to the DDT level in the brain tissue of brown pelicans. The purpose to test the difference between the juveniles and nestings. The following data summary was obtained. (We can assume that the data are normally distributed.)

	Juveniles	Nestings
Sample size	$n_1 = 8$	$n_2 = 6$
sample mean (in ppm)	$\bar{y}_1 = 0.041$	$\bar{y}_2 = 0.022$
sample standard deviation	$s_1 = 0.017$	$s_2 = 0.034$

(i) Can we claim that juveniles have a higher mean concentration? Draw your conclusion with an approximate p-value. (15%)

Sol: We are testing $H_0 : \mu_1 = \mu_2$ against $H_a : \mu_1 \neq \mu_2$. We use the two sample t-test.

$$S_p^2 = \frac{7 \times 0.017^2 + 5 \times 0.034^2}{7 + 5} = 6.5 \times 10^{-4} \text{ or } S_p = 0.0255.$$

By formula

$$T = \frac{0.041 - 0.022}{0.0255\sqrt{1/8 + 1/6}} = 1.38 \sim t_{12}$$

p-value ~ 0.2.

(ii) Suppose the two variances for the two populations are σ_1^2 and σ_2^2 respectively and they are unknown. We also know that

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_2-2}^{n_1-1}.$$

How do we use an F-table to construct a 95% confidence interval for σ_2^2/σ_1^2. Note: $F_{\nu_2,1-a}^{\nu_1} = 1/F_{\nu_1,a}^{\nu_2}$. (10%)

Sol: Ues the F table we have

$$Pr\{\frac{1}{5.29} \leq \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \leq 6.85\} = 0.95.$$

Since $s_2^2/s_1^2 = 4$, the 95% CI for σ_2^2/σ_1^2 is (0.756, 27.4).

6. The natural cure rate of a disease is 20%. A drug is invented and to be tested in a clinical trial with n patients. The clinical trial should satisfied two requirements. (1) The risk (probability) of claiming an ineffective drug as effective should be less than 0.01, and (2) If the drug has cure rate is 30% or higher, we wish to have risk (probability) smaller than 0.05 to miss it. Determine the sample size n for this clinical trial. Note: $Y \sim Bi(n, p) \Rightarrow Y \sim N(np, npq), q = 1 - p$. (20%)

   Sol: We are testing $\text{H}_0 : p \leq 0.2$ against $\text{H}_a : p \geq 0.2$. Let n patients be used and the cured number is Y. Then $RR : Y \geq k$. The requirements are:

$$\begin{aligned} (1)\ p(Y \geq k|p = 0.2) &= 0.01 \Rightarrow 0.01 = P(Z \geq \frac{k - 0.2n}{\sqrt{n \times 0.2 \times 0.8}}) \\ (2)\ p(Y \geq k|p = 0.3) &= 0.95 \Rightarrow 0.95 = P(Z \geq \frac{k - 0.3n}{\sqrt{n \times 0.3 \times 0.7}}) \end{aligned}$$

   They become

$$\begin{aligned} (1)\ k - 0.2n &= 2.33\sqrt{0.16n} \\ (2)\ k - 0.3n &= -1.645\sqrt{0.21n} \end{aligned}$$

   Solve (1) and (2), we get $n = 284$.