Math 341, Final Exam, Spring 2010, Name____________

Student # ____________

Friday, May 14. *Please show complete solution (with all steps) to each problem to receive full credit.*

1. The Rockwell hardness index for steel is determined by pressing a diamond point into the steel and measuring the depth of penetration. For 50 specimens of an alloy of steel, the Rockwell hardness index average 61 with standard deviation 7. The manufacturer claims that this alloy has an average hardness index of at least 63. Is there sufficient evidence to refute the manufacturer's claim? Test at $\alpha = 0.01$ using a one-sided 99% confidence interval. Clearly specify the null and alternate hypotheses and the conclusion. (15 points)

$n = 50 \quad \bar{Y} = 61 \quad S = 7$ $\qquad H_0\ \mu = 63$

$H_1\ \mu < 63$

Reject H_0 if $63 > \bar{Y} + Z_\alpha \frac{S}{\sqrt{n}} = 61 + Z_{.01} \frac{7}{\sqrt{50}}$

$Z_{.01} = 2.33$

$= 61 + \frac{(2.33)(7)}{\sqrt{50}}$

$= 63.30658232$

Do not Reject H_0 because $63 < 63.30658232$

2. The true mean manual dexterity scores of the second graders who participated in sports (μ_1) is being compared with that of second graders who did not participate in sports (μ_2). One wishes to establish that the true mean dexterity score, for second graders who participate is sports, is higher than that of those who did not participate in sports. If the sample sizes for two groups are taken to be the same and the sample variances are $s_1^2 = 4.3$ and $s_2^2 = 4.6$, respectively. Find the sample sizes that give $\alpha = 0.01$ and $\beta = 0.15$, when $\mu_1 - \mu_2 = 2$. (10 points)

$Z_\alpha = Z_{.01} = 2.33$

$Z_\beta = Z_{.15} = \frac{1.04}{3.37}$

$n = \frac{(Z_\alpha + Z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2}$

$\sigma^2 = 4.3 + 4.6 = S_1^2 + S_2^2$

$= 8.9$

$= \frac{(3.37)^2 (8.9)}{2^2}$

$\mu_a = \mu_1 - \mu_2 = 2$

$= \frac{101.07641}{4} = 25.2691025$ $\quad \mu_0 = 0$ (null hypothesis)

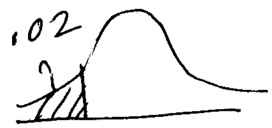
$\mu_1 - \mu_2 = 0$

$n = 26$ //

3. The hourly wages in a particular industry are normally distributed with mean \$13.20 and standard deviation \$3.00. A company in this industry employs 45 workers, paying them an average of \$11.50 per hour. (a) Can this company be accused of paying substandard wages? Use an $\alpha = 0.02$ level test. (b) Compute the type II error probability when $\mu = 13$. (18 points, 9 points each)

$n = 45$
$\bar{Y} = 11.50$

$H_0 \;\; \mu = 13.20$
$H_1 \;\; \mu < 13.20$



$-Z_{.02} = -2.05$

$$Z = \frac{11.50 - 13.20}{3/\sqrt{45}} = -3.80$$

Since $-3.80 < -2.05$ Reject H_0

$$1-\beta = P(Z < -2.05 \mid \mu = 13)$$
$$= P\left(\frac{\bar{Y} - 13.20}{3/\sqrt{45}} < -2.05 \,\Big|\, \mu = 13\right)$$
$$= P\left(\bar{Y} < 13.2 - 2.05\frac{3}{\sqrt{45}} \,\Big|\, \mu = 13\right)$$
$$= P(\bar{Y} < 12.28321213 \mid \mu = 13)$$
$$= P\left(Z < (12.28321213 - 13)/[3/\sqrt{45}]\right)$$
$$= P(Z < -1.60)$$
$$= .0548$$

$$\beta = P(\text{Type II error} \mid \mu = 13)$$
$$= 1 - .0548$$
$$= 0.9452$$

4. The geometric probability mass function is given by $p(y \mid p) = p(1-p)^{y-1}$, y = 1, 2, … A random sample of size n is taken from a population with a geometric distribution. Find the maximum likelihood estimator for p. (10 points)

$$L(p) = p^n (1-p)^{\Sigma y_i - n} \qquad \ln L(p) = n \ln p + (\Sigma y_i - n)\ln(1-p)$$

$$\frac{d}{dp}\ln L(p) = \frac{n}{p} - \frac{(\Sigma y_i - n)}{1-p} = 0 \Rightarrow n(1-p) - p(\Sigma y_i - n) = 0$$

$$n - \cancel{np} - p\Sigma y_i + \cancel{np} = 0 \Rightarrow \hat{p} = \frac{n}{\Sigma y_i}$$

$$\frac{d^2}{dp^2}\ln L(p)\Big|_{p=\hat{p}} = \frac{-n}{\hat{p}^2} - \frac{(\Sigma y_i - n)}{(1-\hat{p})^2} = \frac{-n}{\hat{p}^2} - \frac{(1-\hat{p})}{\hat{p}}\frac{n}{(1-\hat{p})^2} \qquad n = p\Sigma y_i$$

$$= -\frac{n}{\hat{p}^2} - \frac{n}{(1-\hat{p})\hat{p}} < 0, \text{ Hence } \hat{p} = n/\Sigma y_i \text{ is M.L.E.}$$

5. Suppose that $X_1,...,X_n$ and $Y_1,\ldots,Y_n$ are independent random samples form populations with means μ_1 and μ_2 and variances $\infty > \sigma_1^2 > 0$ and $\sigma_2^2 > 0$, respectively. Show that $\bar{X}-\bar{Y}$ is a consistent estimator of $\mu_1 - \mu_2$. (12 points)

Let $W_i = X_i - Y_i$ Then $\bar{X}-\bar{Y} = \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n Y_i}{n}$

$$= \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i}{n} = \frac{\sum_{i=1}^n (X_i - Y_i)}{n} = \bar{W}.$$

$X_i - Y_i = W_i$ forms a random sample because $X_1 \cdots X_n$ & $X_1 \cdots Y_n$ are both random samples independent of eachother and $Var(W_i) = Var(X_i) + Var(Y_i) = \sigma_1^2 + \sigma_2^2 < \infty$. Hence from Law of Large numbers $\bar{W} \xrightarrow{P} EW_1 = \mu_1 - \mu_2$

6. Let $Y_1,...,Y_n$ denote a random sample from the uniform distribution on the interval $(\theta, \theta+1)$. Let $\hat{\theta}_1 = \bar{Y} - \frac{1}{2}$ and $\hat{\theta}_2 = Y_{(n)} - \frac{n}{n+1}$.

   a. Show that $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ.

   b. Find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$. (15 points)

a. $E\hat{\theta}_1 = E\bar{Y} - \frac{1}{2} = \frac{\theta + \theta + 1}{2} - \frac{1}{2} = \theta$

$E\hat{\theta}_2 = [EY_{(n)}] - \frac{n}{n+1}$ —— (*)

$F_{Y_{(n)}}(y) = P(Y_{(n)} \le y) = \left(\frac{y-\theta}{1}\right)^n, \quad \theta \le y \le \theta+1$

$$f_{Y_{(n)}}(y) = \frac{d}{dy} F_{Y_{(n)}}(y) = \begin{cases} n(y-\theta)^{n-1}, & \theta \le y \le \theta+1 \\ 0 & \text{elsewhere} \end{cases}$$

$$E\{Y_{(n)} - \theta\} = \int_\theta^{\theta+1} n(y-\theta)^n\, dy = \frac{n}{n+1}(y-\theta)^{n+1}\Big|_\theta^{\theta+1} = \frac{n}{n+1} \quad \text{—(**)}$$

$$\Rightarrow EY_{(n)} = \theta + \frac{n}{n+1} \underset{\text{from } (*)}{\Rightarrow} E\hat{\theta}_2 = \theta + \frac{n}{n+1} - \frac{n}{n+1} = \theta.$$

(b) $Var(\hat{\theta}_2) = MSE(\hat{\theta}_2) = E(\hat{\theta}_2 - \theta)^2 = E\left(Y_{(n)} - \theta - \frac{n}{n+1}\right)^2$

$$= E(Y_{(n)} - \theta)^2 - \frac{2n}{n+1} E(Y_{(n)} - \theta) + \frac{n^2}{(n+1)^2} \underset{\text{use } (**)}{=} \int_\theta^{\theta+1} n(y-\theta)^{n+1} dy - \frac{2n^2}{(n+1)^2} + \frac{n^2}{(n+1)^2} \quad \varphi$$

$$\varphi \; \frac{n}{n+2}(y-\theta)^{n+2}\Big|_\theta^{\theta+1} - \frac{n^2}{(n+1)^2} = \frac{n}{n+2} - \frac{n^2}{(n+1)^2}.$$

$Var(\hat{\theta}_1) = Var(\bar{Y}) = \frac{Var(Y_1)}{n} = \frac{1}{12n}$ $\quad Var(Y_1) = \frac{1}{12}$.

$$eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{\frac{n}{n+2} - \frac{n^2}{(n+1)^2}}{\frac{1}{12n}} = \frac{12n^2}{(n+2)(n+1)^2}$$

7. Seasonal ranges (in hectors) for alligators were monitored on a lake outside Gainesville, Florida, by biologists from the Florida Game and Fish Commission. Four alligators monitored in the spring showed ranges of 8.0, 12.1, 8.1, and 18.2. Three different alligators monitored in the summer showed ranges of 102.0, 81.7, and 54.7. (a) Estimate the difference between mean spring and summer ranges, with a 99% confidence interval. (b) Are the two means different at level α = 0.01? Why or why not? (c) What statistical assumptions did you make? (20 points)

Spring $n_r = 4$

$\bar{Y}_r = 11.6$

$S_r = 4.796526521$

$S_r^2 = 23.00\dot{6}$

Summer $n_u = 3$

$\bar{Y}_u = 79.4\dot{6}$

$S_u = 23.72895559$

$S_u^2 = 563.06\dot{3}$

99% C.I for

$\mu_u - \mu_r$

(a) $\begin{array}{r}79.4\dot{6} \\ -11.60\end{array} \pm t_{4+3-2,\,.005}\, S_p \sqrt{\frac{1}{4} + \frac{1}{3}}$

$67.8\dot{6} \pm t_{5,\,.005}\,(15.46057351)\sqrt{\frac{1}{4} + \frac{1}{3}}$

$S_p^2 = \dfrac{(23.00\dot{6})3 + (563.06\dot{3})2}{5} = 239.029\dot{3}$

$S_p = 15.46057351$, $t_{5,\,.005} = 4.032$

$67.8\dot{6} \pm 62.3370324\sqrt{.25 + .3\dot{3}}$

$67.8\dot{6} \pm 47.61069492$, $(20.2559174\,,\,115.4773616)$

(b) Yes, because 0 does not belong to $(20.2559174, 115.4773616)$

(c) We assume that the two samples are random samples independent of each other and the population variances are equal. These samples are from normal distributions.