MATH 4720 **TEST 1** **Feb. 10, 2017** **NAME** key

Please show all work for partial credit. Point totals are shown to the left of each problem.

1. The government would like to estimate the proportion of adults that have been contacted by a phishing scheme.
   a) How many adults should be sampled if the government wants their estimate to be within 0.05 of the true proportion with 98% confidence?

(5)

$.05 = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\Rightarrow n = \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{.05}\right)^2$

$n = (.25)\left(\frac{2.326}{.05}\right)^2$

$= 541.03 \rightarrow 542$

   b) If the government takes a random sample of 500 adults, find the probability that the sample proportion will be within 0.05 of the true proportion. Assume the true proportion is 0.40.

(5)

$P(|\hat{p} - .40| < .05)$

$= P(.35 < \hat{p} < .45)$

$z = \frac{.45 - .40}{\sqrt{\frac{(.4)(.6)}{500}}}$

$= 2.28$

$1 - 2 \cdot P(z > 2.28)$

$= .9774$

note: almost 98%

2. Let $Y_1 \ldots Y_n$ be a random sample from $f(y) = \theta y^{\theta-1}, 0 < y < 1$
   Show that $W = \prod_{i=1}^{n} Y_i$ is a sufficient statistic for θ.

(10)

$f(y_1, \ldots y_n; \theta) = \prod_{i=1}^{n} \theta y_i^{\theta-1} = \theta^n \left(\prod_{i=1}^{n} y_i\right)^{\theta-1}$

let $g(u, \theta) = \theta^n \left(\prod_{i=1}^{n} y_i\right)^{\theta-1}$

$h(y_1, \cdots y_n) = 1$

By factorization criterion, $\prod_{i=1}^{n} y_i$, is a sufficient statistic for θ.

3. Let Let $Y_1 \ldots Y_n$ be a random sample from $Y \sim EXP(\theta)$. Find the mean square error for $\hat{\theta} = \frac{n}{n+1}\bar{Y}$

$E(Y) = \theta \quad V(Y) = \theta^2$

(12)

$$E(\hat{\theta}) = \frac{n}{n+1} E(\bar{Y}) = \frac{n}{n+1}\theta \quad \Rightarrow B(\hat{\theta}) = \frac{n}{n+1}\theta - \theta = \theta\left(\frac{-1}{n+1}\right)$$

$$V(\hat{\theta}) = \frac{n^2}{(n+1)^2} V(\bar{Y}) = \frac{n^2\theta^2}{(n+1)^2 \cdot n} = \frac{n\theta^2}{(n+1)^2}$$

$$\Rightarrow MSE(\hat{\theta}) = \frac{n\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} = \frac{(n+1)\theta^2}{(n+1)^2} = \frac{\theta^2}{n+1}$$

4. In a study of the relationship between birth order and college success, an investigator found that 126 in a sample of 180 college graduates were firstborn or only children; in a sample of 100 non-graduates of comparable age and socioeconomic background, the number of firstborn or only children was 54. Estimate the difference in the proportions of firstborn or only children for the two populations from which these samples were drawn. Give a 2σ bound for the error of estimation.

(8)

$$\hat{p}_1 = \frac{126}{180} = .70 \qquad \hat{p}_2 = \frac{54}{100} = .54$$

$$\hat{\sigma}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{(.70)(.30)}{180} + \frac{(.54)(.46)}{100}} = .0604$$

$$2 \cdot \hat{\sigma}_{\hat{p}_1 - \hat{p}_2} = 2(.06) = .12$$

$$(.70 - .54) \pm .12 \rightarrow \boxed{(.04, .28)}$$

5. Let $Y_1 \ldots Y_n$ be a random sample from $Y \sim N(0, \sigma^2)$. Show that $S_n{}^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i{}^2$ is consistent for σ^2.

$$E(Y_i^2) = \sigma^2 + \mu^2 = \sigma^2$$

(12) $$E(S_n^2) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i^2) = \frac{n\sigma^2}{n} = \sigma^2 \Rightarrow S_n^2 \text{ is unbiased}$$

$$\frac{Y_i^2}{\sigma^2} \sim \chi^2_{(1)} \Rightarrow W = \frac{\sum_{i=1}^{n} Y_i^2}{\sigma^2} \sim \chi^2_{(n)}, \quad V(W) = 2n$$

$$V(S_n^2) = V\left(\frac{\sigma^2 \cdot W}{n}\right) = \frac{\sigma^4}{n^2} V(W) = \frac{2\sigma^4}{n}$$

$$\lim_{n\to\infty} V(S_n^2) = \lim_{n\to\infty} \frac{2\sigma^4}{n} = 0 \Rightarrow S_n^2 \xrightarrow{P} \sigma^2$$

6. Let Let $Y_1 \ldots Y_n$ be a random sample from $Y \sim GAM(\alpha = 2, \beta)$.

a) Use the method of moment generating functions to show that $W = {}^{2\sum_{i=1}^{n} Y_i}/_{\beta}$ is a pivotal quantity.

$$m_Y(t) = \frac{1}{(1-\beta t)^2} \qquad m_W(t) = E[e^{tW}] = E\left[e^{t\left(\frac{2\sum Y_i}{\beta}\right)}\right]$$

(5)

$$\overset{\text{ind.}}{=} \prod_{i=1}^{n} E\left[e^{\frac{2tY_i}{\beta}}\right] = \left[m_Y\left(\frac{2t}{\beta}\right)\right]^n = \boxed{\frac{1}{(1-2t)^{2n}} \sim \chi^2_{(4n)}}$$

b) If a sample of size 10 yields a sample mean of $\bar{Y} = 8.6$, find 95% upper confidence interval for β.

(5) $$\frac{2n\cdot\bar{Y}}{\beta} \sim \chi^2_{(4n)}$$

$$P\left[\chi^2_{.95} < \frac{20\cdot\bar{Y}}{\beta}\right] = .95$$

$$\Rightarrow C\left[\beta < \frac{20\cdot\bar{Y}}{\chi^2_{.95}}\right] = .95$$

$$\boxed{C\left[\beta < 6.488\right] = .95}$$

$n = 10$

$$\Rightarrow \frac{20\cdot\bar{Y}}{\beta} \sim \chi^2_{(40)}$$

$$C\left[\beta < \frac{20(8.6)}{26.5093}\right] = .95$$

7. A city health department wishes to estimate the mean bacteria count per unit volume of water at a lake beach. A researcher collected 50 water samples of unit volume and found the average bacteria count to be 193.7 with a standard deviation of 21.0.

   a) Find a 99% lower confidence interval for the true mean bacteria count per unit volume at the lake beach.

(8)

$C\left[\mu > \bar{y} - z_{.01} \cdot s/\sqrt{n}\right] = .99$

$C\left[\mu > 193.7 - 2.326\,(21)/\sqrt{50}\right] = .99$

$\boxed{C\left[\mu > 186.79\right] = .99}$

   b) The beach manager claims the average bacteria count is below 180. Use your answer to part (a) to support or refute the manager's claim.

(2) Refute. We are highly confident (99%) that μ is at least 186.79

8. A researcher would like to compare two different brands of pain medication. The sample data (in minutes) for both brands is listed below:

Brand	sample size	mean relief time	standard deviation of relief time
A	15	44	11
B	12	53	9

Construct a 98% confidence interval for the difference in mean relief times for the two brands of pain medication. Assume relief times are approximately normally distributed. Is there a significant difference in the two brands of pain medications?

(8)

$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$s_p^2 = \frac{14(11^2) + 11(9^2)}{15+12-2} = 103.4$

d.f. $= 25$ $\quad t_{25,.01} = 2.485$

$(44-53) \pm 2.485\sqrt{103.4}\sqrt{\frac{1}{15} + \frac{1}{12}}$

-9 ± 9.7866

$\boxed{(-18.7866, .7866)}$

no significant difference in pain medications

(0 in CI)

9. UVU is experimenting with hybrid courses for MATH 1050, College Algebra. The math department is concerned that in hybrid courses, the variability of grades will increase. With conventional instruction, the variance is $\sigma^2 = 400$. A random sample of 51 students who took MATH 1050 as a hybrid was collected and the average grade was $\overline{y} = 77$ with a standard deviation of $s = 24.5$. Assume grades are normally distributed. Find a 90% lower confidence bound for σ^2 for the hybrid courses. Should the math department be concerned?

(8)

$$C\left[\sigma^2 > \frac{(n-1)s^2}{\chi^2_{.10}}\right] = .90$$

$$C\left[\sigma^2 > \frac{50(24.5)^2}{63.1671}\right] = .90$$

$$\boxed{C[\sigma^2 > 475] = .90}$$

Yes, the math dept. should be concerned. We are 90% confident that σ^2 is at least 475 ($\sigma^2 > 400$).

10. Let $Y_1 \ldots Y_n$ be a random sample from $f(y) = \frac{2y}{\theta^2}, 0 < y < \theta$

a) Show that $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$ and $\hat{\theta}_2 = Y_1 + \frac{1}{2}Y_2$ are unbiased.

(4)

$$E(Y) = \int_0^\theta \frac{2y^2}{\theta^2}\,dy = \frac{2\theta}{3}$$

$$E(\hat{\theta}_1) = \frac{3}{2}E(\overline{Y}) = \frac{3}{2}\left(\frac{2\theta}{3}\right) = \boxed{\theta}$$

$$E(\hat{\theta}_2) = E(Y_1) + \frac{1}{2}E(Y_2)$$

$$= \frac{2}{3}\theta + \frac{1}{2}\left(\frac{2}{3}\theta\right) = \boxed{\theta}$$

b) Find the relative efficiency of $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$ to $\hat{\theta}_2 = Y_1 + \frac{1}{2}Y_2$.

(8)

$$E(Y^2) = \int_0^\theta \frac{2y^3}{\theta^2}\,dy = \frac{2y^4}{4\theta^2}\Bigg|_0^\theta = \frac{\theta^2}{2}$$

$$\Rightarrow V(Y) = \frac{\theta^2}{2} - \frac{4\theta^2}{9} = \frac{\theta^2}{18}$$

$$V(\hat{\theta}_1) = \frac{9}{4}V(\overline{Y}) = \frac{9}{4}\cdot\frac{\theta^2}{18n} = \frac{\theta^2}{8n}$$

$$V(\hat{\theta}_2) = V(Y_1) + \frac{1}{4}V(Y_2) = \frac{\theta^2}{18} + \frac{\theta^2}{72} = \frac{5\theta^2}{72}$$

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\frac{5\theta^2}{72}}{\frac{\theta^2}{8n}} = \frac{5(8n)}{72} = \boxed{\frac{40n}{72}}$$

note: > 1 for $n \geq 2$

i.e. $\hat{\theta}_1$ better than $\hat{\theta}_2$