UNIVERSITY OF CALGARY

DEPARTMENT OF MATHEMATICS AND STATISTICS

STAT 327 — Test no. 2 — April 3, 2009

Marks

NAME: Solutions

10 3 1. (a) In order to estimate the proportion p of people in a large city who favor a proposed tax increase, it was decided to conduct a survey of n randomly chosen people in the city. How large should n be in order that the resulting margin of error in a 95% confidence interval for p will be no larger than 0.04?

$$n \geq \frac{(1.96)^2 \, 1/4}{(.04)^2} = 600.25 \quad \text{so } n = 601$$

2 1

4 (b) Suppose that a random sample of 800 people yields the result that 350 of them favor the tax increase. Find a 95% confidence interval for p.

$$\hat{p} = \frac{350}{800} = 0.4375$$

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\underbrace{.4375}_{1} \pm \underbrace{1.96\sqrt{\frac{(.4375)(.5625)}{800}}}_{2}$$

$$.4375 \pm 0.0344$$

$$0.4031 < p < 0.4719 \quad \leftarrow 1$$

3 (c) Use the data given in part (b), find the P-value for testing the null hypothesis $p = 0.5$ against the alternative hypothesis $p < 0.5$.

$$H_0: p = 0.5$$

$$H_a: p < 0.5$$

$$z = \frac{\hat{p} - 0.5}{\sqrt{(.5)(.5)/800}} = -3.54 \leftarrow 2$$

$$p\text{-value} = P[Z < -3.54] \approx .000232$$

1

10 2. Assume that the distribution of family incomes in a city is approximately normally distributed, with an unknown population mean and an unknown population standard deviation. A random sample of seven families from the city yields the following incomes (in thousands of dollars):

$$20, 90, 60, 30, 70, 40, 80.$$

5 (a) Compute point estimates of the population mean and the population standard deviation.

$$\bar{x} = \frac{20+90+60+30+70+40+80}{7} = \underline{55.714} \text{ (thousands of \$)}$$

2

$$\sum x_i^2 = 400+8100+3600+900+4900+1600+6400 = 25900$$

$$\text{so } s^2 = \frac{25900-(390)^2/7}{6} = 695.238$$

$$s = \underline{26.36} \text{ (thousands of \$)}$$

3

5 (b) Carry out a test of the null hypothesis that the population mean income is \$50,000 against the alternative hypothesis that it is greater than \$50,000. Use significance level $\alpha = 0.10$.

$H_0: \mu = 50$ (thousands of \$) $\quad \alpha = .10$

$H_1: \mu > 50$

Reject H_0 if $z = \frac{\bar{x}-50}{s/\sqrt{7}} > t_{6,.10} = 1.440$ 1

3

$$z = \frac{5.714\sqrt{7}}{26.36} = 0.573 < 1.440$$

so $\underline{\text{accept } H_0}$ or there is not enough evidence to reject H_0 at level $\alpha = 0.10$.

1

10 3. A random sample of 400 women revealed that 72 smoked at least one pack of cigarettes per day. Also, a random sample of 500 men revealed that 70 smoked at least one pack of cigarettes per day.

5 (a) Find a 90% confidence interval for the difference between the proportion of women and the proportion of men who smoke at least one pack per day.

pop. 1 = women, pop. 2 = men (say)

Then $\hat{p}_1 = \frac{72}{400} = 0.18$, $\hat{p}_2 = \frac{70}{500} = 0.14$

$$\hat{p}_1 - \hat{p}_2 \pm 1.645 \sqrt{\frac{(.18)(.82)}{400} + \frac{(.14)(.86)}{500}}$$

$$0.04 \pm 1.645\,(.02469)$$

$$.04 \pm .0406$$

$$-.0006 < p_1 - p_2 < .0806$$

5 (b) For testing the null hypothesis that the two proportions are the same against the alternative hypothesis that the two proportions are different, find the P-value of the data.

$$H_0: p_1 = p_2$$

$$H_a: p_1 \neq p_2$$

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

pooled est $\hat{p}$ under H_0 is

$$\hat{p} = \frac{72+70}{400+500} = 0.1578$$

$$= \frac{.04}{\sqrt{.1578\left(\frac{1}{400} + \frac{1}{500}\right)}} = 1.50$$

$$P\text{-value} = 2P[Z > 1.50] = 2(.0668) = .1336$$

10 4. A coffee manufacturer is interested in whether the mean daily consumption of regular-coffee drinkers is less than that of decaffeinated-coffee drinkers. A random sample of 50 regular-coffee drinkers showed a sample mean of 4.35 cups per day, with a sample standard deviation of 1.20. A random sample of 40 decaffeinated-coffee drinkers showed a sample mean of 5.12 cups per day, with a sample standard deviation of 1.36.

5 (a) Set up the appropriate null and alternative hypotheses (based on the first sentence above) and carry out the test to obtain a conclusion at the 0.01 significance level.

$H_0: \mu_R = \mu_D \quad (\mu_R - \mu_D = 0) \qquad \alpha = .01$

$H_a: \mu_R < \mu_{Dec} \quad (\mu_R - \mu_D < 0)$

Reject H_0 if $Z = \dfrac{\bar{X}_R - \bar{X}_D}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} < -Z_{.01} = -2.326$

$$Z = \frac{4.35 - 5.12}{\sqrt{\frac{(1.20)^2}{50} + \frac{(1.36)^2}{40}}} = \frac{-0.77}{0.2737} = \underline{-2.81}$$

3

$< \underbrace{-2.326}_{1}$, $\underbrace{\text{so reject } H_0 \text{ at level } \alpha = .01}_{1}$

5 (b) For the above hypothesis test, compute the P-value of the data.

$$\text{P-value} = \underbrace{P[Z < -2.81]}_{3} = .0025 \quad 2$$

40 ← TOTAL MARKS