Student Number:_____ **Final Exam ANSWERS**
PBAF 527, Winter 2005 **Due March 14, noon**

Ground rules:

- Breathe!
- DO NOT WRITE YOUR NAME ON THE EXAM—write your student number. If you're worried about losing the first page, write your student number on each page.
- You can use your notes or books, a calculator or a spreadsheet, but you may not communicate with other people about this midterm nor the material covered by it.
- In order to receive as much credit as possible, please show all of your work. Showing that you understand the question and know how to set up the solution correctly is more important than arriving at the exact answer.
- Read each question carefully and answer all parts of each question.
- Good luck!
- When you are done, please put it under my door (209C Parrington) or email it to me. If you put it under my door, please email me to let me know you have done so.
- Please note I have changed the time the final is due to noon on Monday, March 14 (from 9am as I said in my original email).

Which is more important, parental involvement or the cost of cigarettes in reducing teen smoking? Given the established evidence of the detrimental impact of smoking on health, policy makers trying to reduce youth smoking through a variety of ad campaigns, clean indoor air laws, taxes, youth access restrictions, and advertising restrictions. Recent efforts have focused on soliciting parental aid in reducing smoking among youth. How successful is such a policy likely to be?

A 1996 survey of high school students in the United States interviewed a representative sample of 11,237 high school students, including 3,111 smokers and 8,125 nonsmokers. The students were asked whether they currently smoked and about their interactions with their parents. 28% of students smoked at the time of the study. Among smokers, the percent who spoke with their parents about daily issues was 49%, while among non-smokers 52% spoke with their parents that frequently about daily issues.

The researchers also figured out the average cost of cigarettes for both smokes and non-smokers, based upon their state of residence. For non-smokers, the average cost was $1.89 (with a standard deviation of $0.23) and for smokers it was 1.87 (with a standard deviation of $0.21).

As you answer the following questions, make sure to state any assumptions you must make. Use a=0.01.

Summary of data provided:

Group	n	Parent Interaction	Mean Cigarette Price	s of Cigarette Price
1. Smokers	3111	.49	1.87	.21
2. Non-Smokers	8125	.52	1.89	.23

in 2004, 22% of the youth population smoked
in 1996, 28% of the sample smoked (3,111/11,237)

1. Last summer, the Centers for Disease Control (CDC) reported that teen smoking had fallen to the lowest levels since the CDC started to keep track in 1975. They reported that in June 2004 that the teen smoking rate was 22%. Was the rate of smoking in 1996 really higher than in 2004? Make sure you formally set up your hypothesis, carry out all the steps, and interpret your findings. (10 points)

THIS QUESTION WAS MISFORMULATED IN YOUR COPY OF YOUR EXAM. SO, IT NOW COUNTS AS EXTRA CREDIT. ANYTHING YOU DID WELL COUNTS IN YOUR FAVOR.

H_0: p=.22
H_a: p>.22

a=.01
one-sided test, so z_a=2.33
decision rule: if $z>z_a$ then we can reject the null hypothesis.

test statistic $z = \dfrac{\hat{p} - p_0}{s_{\hat{p}}} = \dfrac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \dfrac{.28 - .22}{\sqrt{.22(1-.22)/11{,}237}} = 15.35$

Since 15.35>2.33 we can reject the null hypothesis at a .01 level. The rate of smoking among teens was really higher in 1996.

2. Create and interpret a range estimate for the proportion of youth who smoked overall in 1996. (10 points)

$\hat{p} \pm z_{a/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = .28 \pm 2.58\sqrt{\dfrac{.28(1-.28)}{11{,}237}} = .28 \pm .01$

P(.27<*p*<.29)=.99

We are 99% confident that the true rate of youth smoking in 1996 was between 27% and 29%.

3. Is parental interaction different for smokers and nonsmokers? Formally set up, perform, and interpret a hypothesis test. (10 points)

Assumption: 2 independent sample; sample size is large enough that the normal distribution is an adequate approximation for the sampling distribution of $\hat{p}_1$, $\hat{p}_2$, and $\hat{p}_1 - \hat{p}_2$

H_0: $p_1-p_2=0$
H_a: $p_1-p_2 ? 0$
where p_1 is the true proportion of youth nonsmokers, and p_2 is the true proportion of youth smokers.
a=.01
two-sided test, so $z_{a/2}=2.58$
decision rule: if $|z|>z_{a/2}$ then we can reject the null hypothesis.
Test statistic:

First find se:

$$s_{(\hat{p}_1-\hat{p}_2)} = \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{.51(1-.51)\left(\frac{1}{8125}+\frac{1}{3111}\right)} = \sqrt{0.2498633(0.000444517)} = 0.0105$$

where $\hat{p} = \dfrac{x_1+x_2}{n_1+n_2} = \dfrac{4225+1524}{8125+3111} = 0.51$ $\quad x_1 = .52(8125) = 4225$ $\quad x_2 = .49(3111) = 1524$

$$z = \frac{\hat{p}_1-\hat{p}_2}{s_{(\hat{p}_1-\hat{p}_2)}} = \frac{.52-.49}{0.0105} = 2.85$$

Because 2.85>2.58 we can reject the null hypothesis at a .01 level of significance. There is a difference in interaction between teen smokers and nonsmokers and their parents.

4. Are cigarette prices different for smokers and nonsmokers? (Remember, the researchers figured out the average cost of cigarettes in the state of each person in their sample.) Formally set up, perform, and interpret a hypothesis test. (10 points)

H_0: $\mu_1 - \mu_2 = 0$
H_a: $\mu_1 - \mu_2 ? 0$
where μ_1 is the true proportion of youth nonsmokers, and μ_2 is the true proportion of youth smokers.
a=.01
two-sided test, so $z_{a/2}=2.58$
decision rule: if $|z|>z_{a/2}$ then we can reject the null hypothesis.
Test statistic:

First find se:

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = \sqrt{\frac{.23^2}{8125}+\frac{.21^2}{3111}}$$

$$= \sqrt{\frac{.0529}{8125}+\frac{.0441}{3111}} = \sqrt{.000006512+.000014176} = \sqrt{.00002069} = 0.004548$$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{s_{(\hat{p}_1-\hat{p}_2)}} = \frac{1.89-1.87}{0.004548} = 4.397$$

Because 4.40>2.58 we can reject the null hypothesis at a .01 level of significance. There is a difference in the cost of cigarettes for teen smokers and nonsmokers.

5. How strong is your evidence in number 3? How about number 4? For which is the evidence stronger? (10 points)

We want the 2-sided p-values associated with calculated t-statistics.

P(z<-2.85 and z>2.85)=1-(2*.4978)=.004 or 2(.5-.4978)=2(.0022)=.0044. (if I look up 2.85 in the normal table, I get a probability from 0 to 2.85 of .4978. Since I want the tails on either end, I need to multiply the probability by 2 and subtract it from 1).

P(z<-4.40 and z>4.40)= something very small. The value exceeds those in the normal table, meaning that the probability is smaller than that for ±2.85, less than .001 (it is actually closer to 0.00001, that is less than 0.001%--really, really small).

Because of the smaller p-value, we have stronger evidence for the difference in the price of cigarettes than for interactions with parents, although both certainly have strong evidence for the differences.

6. Create a range estimate for the difference in the proportion of smokers and non-smokers who spoke to their parents daily about everyday issues. (10 points)

$$(\hat{p}_1-\hat{p}_2)\pm z_{a/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}\approx(\hat{p}_1-\hat{p}_2)\pm z_{a/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$=(.52-.49)\pm 2.58\sqrt{\frac{.52(1-.52)}{8125}+\frac{.49(1-.49)}{3111}}=.03\pm 2.58\sqrt{\frac{.2496}{8125}+\frac{.2499}{3111}}$$

$$=.03\pm 2.58\sqrt{.00003072+.0000803}=.03\pm 2.58\sqrt{.0003072+.0000803}$$

$$=.03\pm 2.58\sqrt{.00011}$$

$$=.03\pm 2.58(0.0105)$$

$$=.03\pm 0.027$$

$P(.003<p_1-p_2<.06)=.99$

We are 99% confident that the proportion of nonsmokers speaking to their parents daily about everyday issues is between .3 percentage points and 6 percentage points more than the proportion of smokers who speak to their parents daily about everyday issues.

7. Create a range estimate for the difference in the price of cigarettes for smokers and non-smokers. (10 points)

$$(\bar{x}_1-\bar{x}_2)\pm z_{a/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\approx(\bar{x}_1-\bar{x}_2)\pm z_{a/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$=(1.89-1.87)\pm 2.58\sqrt{\frac{.23^2}{8125}+\frac{.21^2}{3111}}$$

$$=.02\pm 2.58\sqrt{\frac{.0529}{8125}+\frac{.0441}{3111}}=.02\pm 2.58\sqrt{.0000065+.0000142}$$

$$=.02\pm 2.58\sqrt{.0000207}$$

$$=.02\pm 2.58(0.004548)$$

$$=.02\pm .01$$

$P(.01<\mu_1-\mu_2<.03)=.99$

So, we are 99% confident that the price of cigarettes that nonsmokers face is 1 cent to 3 cents more per package than the price of cigarettes that smokers face in their states.

8. Write an essay, summarizing and interpreting ALL your findings for a non-technical policy audience. In your response, make sure to (1) answer the question of which is more effective in reducing teen smoking, parental involvement or the cost of cigarettes and (2) discuss the difference between statistical differences and meaningful differences when conducting hypothesis tests. (20 points)

   ONE POSSIBLE RESPONSE (BASED ON STUDENTS RESPONSES):

   With substantial medical evidence known today on the dangers of tobacco use, smoking among America's youth is an important health and social issue. Good news is available, however, indicating that the trend of teenage smoking is decreasing. A study of smoking among 11,237 randomly selected American high school students conducted in 1996 found a rate of smoking among respondents to be 28%, and estimated that between 27% and 29% of U.S. teens smoked.[1] Eight years later, the CDC reported that the rate of teen smoking in the U.S. was 22%--this was a real decline in the smoking rate among youth.[2]

   Numerous factors can contribute to the youth smoking rate and its reduction. The 1996 study evaluated two aspects of teenage smoking: interaction with parents and the average cost of cigarettes. Of responding high school students who smoked, 49% spoke with their parents about everyday issues; among nonsmoking respondents, 52% did the same. This 3% difference in parental interaction was large enough to be a non-random difference.[3] The study convincingly showed that smokers and non-smokers experience different levels of parental involvement.[4] In fact, we estimate that nonsmokers interact with their parents about daily issues between .3 and 6 percentage points more than do smokers.[5] And stronger evidence showed that the price of cigarettes available to the average smoker is between $.01 and $.03 less than to the average non-smoker, based on their state of residence.[6]

   However, it is not clear from this study which factor may be more effective in reducing teen smoking, if at all. Hypothesis testing only allows us to confidently state that the differences witnessed in the survey are real differences between smokers and non-smokers. But the statistical differences do not tell us whether the differences are large enough to be of value in a practical sense. For instance, the 1996 study showed a real difference in the cost of cigarettes for smokers than for non-smokers, based on their state of residence. But the cost difference is only a few pennies. Is this a meaningful difference that would affect the rate of youth smoking? Furthermore, we know from these data that smokers and nonsmokers are different, not how either cigarette prices or direct parental intervention reduce smoking.

---

[1] 99% confidence interval.
[2] Hypothesis test of one mean, significant at a .01 level, t=15.35.
[3] Large sample hypothesis test of two proportions, significant at a .01 level. t=2.85.
[4] p=.004
[5] 99% confidence interval.
[6] 99% confidence interval; hypothesis test of two means, unequal variances assumed, significant at a .01 level; p=.002.