Statistics 1040, Sections 002, 003 & 004, Quiz 12 (20+ Points)

April 26, 2002

Your Name: ____________________

You have exactly 15 minutes for this quiz. Each of the three questions is worth 20 points. Choose one question as your "main question" and do it completely. Clearly indicate which question you have chosen. You may also earn extra credit if you manage to work on more than one questions during this time.

Question 1: from: Stat 1040, Fall 2000, Final Test, Thursday, December 14, 2000 (Question 9)

(20 Points) The spermicide nonoxynol–9 kills HIV in the test tube, so researchers hypothesized that it might be useful in protecting high–risk women from HIV. Other researchers argued that nonoxynol–9 might increase the risk because it is an irritant. In a study of 990 prostitutes, participants were randomly divided into two groups. The treatment group were given a nonoxynol–9 gel. The control group were given a similar–looking but inactive gel. When the study ended in May 2000, 67 of the 495 women in the treatment group were HIV–positive, and 44 of the 495 women in the control group were HIV–positive. Perform a 2–tailed test to decide whether the treatment and control groups were significantly different with respect to HIV. Clearly state your conclusions.

2-sample z-test

1. Null: treatment makes no difference w.r.t. HIV, i.e., $\%_{treat} = \%_{control}$

   Alternative: treatment does make a difference, i.e., $\%_{treat} \neq \%_{control}$

2. Treatment

   $SD = \sqrt{\frac{67}{495} \cdot \frac{428}{495}} = 0.342$

   $\%_{treat} = \frac{67}{495} = 13.5\%$

   $SE_{sum_{treat}} = \sqrt{495} \cdot 0.342 = 7.6$

   $SE_{\%_{treat}} = \frac{7.6}{495} \cdot 100\% = 1.5\%$

   Control

   $SD = \sqrt{\frac{44}{495} \cdot \frac{451}{495}} = 0.285$

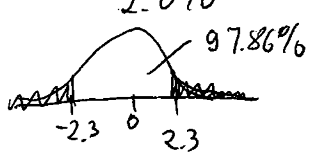
   $\%_{control} = \frac{44}{495} = 8.9\%$

   $SE_{sum_{control}} = \sqrt{495} \cdot 0.285 = 6.3$

   $SE_{\%_{control}} = \frac{6.3}{495} \cdot 100\% = 1.3\%$

   $SE_{diff} = \sqrt{1.5^2 + 1.3^2} = 2.0\%$

   $z = \frac{13.5\% - 8.9\%}{2.0\%} = 2.3$

3. 

   P-value = area outside = 100% − 97.86% = 2.14%

4. reject the null hypothesis (P-value < 5%)
   - result statistically significant
   - treatment makes a difference w.r.t. HIV (i.e., nonoxynol-9 increases the risk of HIV)

Question 2: from: Stat 1040, Spring 2000, Final Test, Friday, May 5, 2000 (Question 5)

(20 Points) Reading test scores are known to follow the normal curve. The average reading test score is supposed to be 100. I suspect that the children at one of the local schools have a higher average reading score, so I take a simple random sample of 10 of these children and find that their average reading score is 112, with an SD of 18. Is my suspicion correct? Perform the appropriate statistical test. You must clearly state a null and alternative hypothesis, compute a test statistic and a P-value and state your conclusions.

sample size = 10 (< 30), assume "normal" & SD box unknown → t-test

① Null: same score as elsewhere, i.e., box avg (for all children at this school) = 100

Alternative: higher score as elsewhere, i.e., box avg > 100

② $SD^{+} = \sqrt{\frac{10}{10-1}} \cdot 18 = 19$

$SE_{sum} = \sqrt{10} \cdot 19 = 60$

$SE_{avg} = \frac{60}{10} = 6$

$t = \frac{112-100}{6} = 2$ $\quad df = 10 - 1 = 9$

③ in table: 2.26 ⇝ 2.5%, 1.83 ⇝ 5%

so t = 2 ⇝ p-value between 2.5% and 5%

④ reject the null hypothesis (p-value < 5%)

- result statistically significant
- students at this school have a higher average score than students elsewhere

Question 3: from: Stat 1040, Spring 1999, Final Test, Monday, May 3, 1999 (Question 8)

(20 Points) In the game of chess, the first few moves play a very important role in determining the final outcome. Five different opening strategies are highly favored by chess experts. To determine whether one or more of these strategies is most preferred by grand masters in international competition, a random sample of 100 grand masters is taken, and each is asked which of the strategies he or she would prefer to employ. A summary of their responses is given below:

Strategy	A	B	C	D	E
Frequency	17	27	22	15	19

Make a χ^2–test of the null hypothesis that there is no preference between these strategies by grand masters in international competition. You should state the null and the alternative hypotheses and clearly state your conclusions.

① Null: no preference – choosing at random, chance for each strategy is 1/5

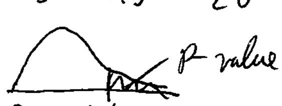
Alternative: some strategies are preferred, chances not all equal to 1/5

②

Strategy	Obs	Exp	$\frac{(Obs-Exp)^2}{Exp}$
A	17	20	0.45
B	27	20	2.45
C	22	20	0.20
D	15	20	1.25
E	19	20	0.05

sum → $\chi^2 = 4.4$

df = 5 − 1 = 4

③ 

P-value between 30% and 50%
($\chi^2 = 4.4$ between 4.88 and 3.36)

④ do not reject the null hypothesis (P-value > 5%)

- no preference – strategies chosen at random

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of }[(\text{deviations from box average})^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

$$\text{EV}_{avg} = \text{box average}$$

$$\text{SE}_{avg} = \frac{\text{SE}_{sum}}{\text{number of draws}}$$

$$\text{SD}_{0\text{-}1\text{ box}} = \sqrt{\begin{matrix}\text{fraction} \\ \text{of } \boxed{1}\text{'s}\end{matrix} \times \begin{matrix}\text{fraction} \\ \text{of } \boxed{0}\text{'s}\end{matrix}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1}\text{'s in the box}$$

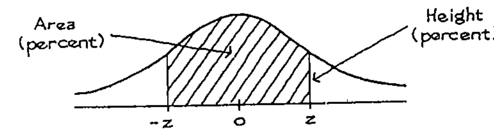
$$\text{SE}_{\%} = \frac{\text{SE}_{\text{sum}}}{\text{number of draws}} \times 100\%$$

$$\text{SD}^{+} = \sqrt{\frac{\text{number of draws}}{\text{number of draws} - 1}} \times \text{SD}$$

$$\text{SE}_{\text{diff}} = \sqrt{a^2 + b^2},$$

where a is the SE for the first quantity, b is the SE for the second quantity, and the two quantities are independent

Tables

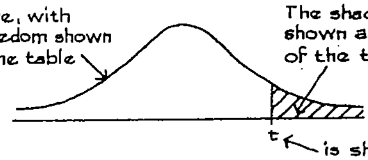


A NORMAL TABLE

z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	Q1 2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

A *t*-TABLE

Student's curve, with degrees of freedom shown at the left of the table



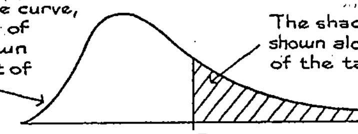
The shaded area is shown along the top of the table

t is shown in the body of the table

Degrees of freedom	25%	10%	5%	2.5%	1%	0.5%
1	1.00	3.08	6.31	12.71	31.82	63.66
2	0.82	1.89	2.92	4.30	6.96	9.92
3	0.76	1.64	2.35	3.18	4.54	5.84
4	0.74	1.53	2.13	2.78	3.75	4.60
5	0.73	1.48	2.02	2.57	3.36	4.03
6	0.72	1.44	1.94	2.45	3.14	3.71
7	0.71	1.41	1.89	2.36	3.00	3.50
8	0.71	1.40	1.86	2.31	2.90	3.36
Q2 → 9	0.70	1.38	1.83	2.26	2.82	3.25
10	0.70	1.37	1.81	2.23	2.76	3.17
11	0.70	1.36	1.80	2.20	2.72	3.11
12	0.70	1.36	1.78	2.18	2.68	3.05
13	0.69	1.35	1.77	2.16	2.65	3.01
14	0.69	1.35	1.76	2.14	2.62	2.98
15	0.69	1.34	1.75	2.13	2.60	2.95
16	0.69	1.34	1.75	2.12	2.58	2.92
17	0.69	1.33	1.74	2.11	2.57	2.90
18	0.69	1.33	1.73	2.10	2.55	2.88
19	0.69	1.33	1.73	2.09	2.54	2.86
20	0.69	1.33	1.72	2.09	2.53	2.85
21	0.69	1.32	1.72	2.08	2.52	2.83
22	0.69	1.32	1.72	2.07	2.51	2.82
23	0.69	1.32	1.71	2.07	2.50	2.81
24	0.68	1.32	1.71	2.06	2.49	2.80
25	0.68	1.32	1.71	2.06	2.49	2.79

A CHI-SQUARE TABLE

The chi-square curve, with degrees of freedom shown along the left of the table



The shaded area is shown along the top of the table

is shown in the body of the table

Degrees of freedom	99%	95%	90%	70%	50%	30%	10%	5%	1%
1	0.00016	0.0039	0.016	0.15	0.46	1.07	2.71	3.84	6.64
2	0.020	0.10	0.21	0.71	1.39	2.41	4.60	5.99	9.21
3	0.12	0.35	0.58	1.42	2.37	3.67	6.25	7.82	11.34
Q3 → 4	0.30	0.71	1.06	2.20	3.36	4.88	7.78	9.49	13.28
5	0.55	1.14	1.61	3.00	4.35	6.06	9.24	11.07	15.09
6	0.87	1.64	2.20	3.83	5.35	7.23	10.65	12.59	16.81
7	1.24	2.17	2.83	4.67	6.35	8.38	12.02	14.07	18.48
8	1.65	2.73	3.49	5.53	7.34	9.52	13.36	15.51	20.09
9	2.09	3.33	4.17	6.39	8.34	10.66	14.68	16.92	21.67
10	2.56	3.94	4.86	7.27	9.34	11.78	15.99	18.31	23.21
11	3.05	4.58	5.58	8.15	10.34	12.90	17.28	19.68	24.73
12	3.57	5.23	6.30	9.03	11.34	14.01	18.55	21.03	26.22
13	4.11	5.89	7.04	9.93	12.34	15.12	19.81	22.36	27.69
14	4.66	6.57	7.79	10.82	13.34	16.22	21.06	23.69	29.14
15	5.23	7.26	8.55	11.72	14.34	17.32	22.31	25.00	30.58
16	5.81	7.96	9.31	12.62	15.34	18.42	23.54	26.30	32.00
17	6.41	8.67	10.09	13.53	16.34	19.51	24.77	27.59	33.41
18	7.00	9.39	10.87	14.44	17.34	20.60	25.99	28.87	34.81
19	7.63	10.12	11.65	15.35	18.34	21.69	27.20	30.14	36.19
20	8.26	10.85	12.44	16.27	19.34	22.78	28.41	31.41	37.57

Source: Adapted from p. 112 of Sir R. A. Fisher, *Statistical Methods for Research Workers* (Edinburgh: Oliver & Boyd, 1958).