9.07 Practice Final Solutions – Fall 2004

1. Health and Nutrition Examination Survey of 1976-1980 recorded handedness of the 2,237 surveyed Americans aged 25-34:
   At 1% significance level can you conclude that gender and handedness are not I independent?

	Right Handed	Left Handed	Can be either
Male	934	113	20
Female	1,070	92	8

Solution:
Use χ^2 test for independence. Compute total counts:

	Right Handed	Left Handed	Can be either	Total
Male	934	113	20	1,067
Female	1,070	92	8	1,170
Total	2,004	205	28	2,237

Compute Expected counts in each cell:

	Right Handed	Left Handed	Can be either
Male	2,004x1,067/2,237=956	205x1,067/2,237=98	28x1,067/2,237=13
Female	2,004x1,070/2,237=1,048	205x1,070/2,237=107	28x1,070/2,237=15

Thus $\chi^2_{obt} = (934\text{-}956)^2/956 + (113\text{-}98)^2/98 + (20\text{-}13)^2/13 +$
$(1{,}070\text{-}1{,}048)^2/1{,}048 + (92\text{-}107)^2/107 + (8\text{-}15)^2/15 = 12$

$\chi^2_{crit} = 9.210$ for $\alpha=0.01$ and df=(3-1)(2-1)=2

Since $\chi^2_{obt} = 12 > 9.210 = \chi^2_{crit,}$ we reject H0 and conclude that gender and handedness are not independent.

2. A study published by the government's Uniformed Services University of the Health Sciences in Bethesda, Md monitored the effect of taking aspirin on cardiac patients. 80%(751) of the 936 study participants were given aspirin. During the 2 years of the study, 22 of the 936 patients died from a heart attack. Of these, 10 were among the 185 who had not had aspirin to prevent a heart attack. The study concluded that fatality for non-users of aspirin was triple that for those on daily aspirin dose. At 1% significance level, test if the fatality rate of the aspirin non-users is twice that for those on daily aspirin dose.

Solution:
Let X=deaths among aspirin non-users, Y=deaths among aspirin users.
Then X ~ Bin(185, p_X) and Y ~ Bin(751, p_Y).

Want to test H_0: $p_X=2p_Y$ vs H_1: $p_X>2p_Y$.
Use Normal approximation to the Binomial:
$Zobt=(2\,p_Y' - p_X')/\sqrt{4\,p_Y'(1-p_Y')/751 + p_X'(1-p_X')/185}=$
$(2 \times 0.016 - 0.054)/\sqrt{4 \times 0.016 \times (1-0.016)/751 + 0.054(1-0.054)/185}=$
= -1.16. p-value=P(Z<zobt)=.1230. Comparing p-value with 0.01 significance level, we retain H0: no significant evidence to conclude that fatality from heart attacks in cardiac patients on aspirin is twice lower than that of non-aspirin users.

3. Answer the following quickies with an answer and a short explanation
   a) The chance of flipping exactly 50 heads in 100 independent tosses of a fair coin is nearly half. True or False? Explain.

   Solution: False.
   X=# heads in 100 coin tosses of a fair coin ~ Bin(100, p=1/2)
   P(X=50) = 100!/(50!50!) (.5)^50 (.5)^50 <<1/2

   Intuitively, 49 and 51 heads are almost as likely as 50, so there must be a less than 50% chance of getting 50.

   b) If X denotes the number of successes in n independent Bernoulli trials, each trial having success probability p, and if Y denotes the number of failures, what is the variance of X-Y?

   Note that Y=n-X; Var(X-Y)=Var(X-(n-X))=Var(2X-n)
   =Var(2X)=4Var(X)=4np(1-p)

   c) What is the variance of X+Y (from part b)?
   X+Y=n = const; Var(X+Y)=Var(n)=0

   d) When estimating a proportion, a random sample of size 200 from a population of 20,000 is as accurate as a random sample of size 400 from a population of 40,000, assuming the true proportions are the same for both populations. True or False? Explain.

   False. SE(p1')=sqrt(p(1-p)/200) > SE(p2')=sqrt(p(1-p)/400), thus estimate with a sample of size 400 will be more accurate.

   e) Using a t-test instead of a z-test to estimate a mean when a standard deviation is known, the data is roughly normal, and the sample size is small (under 30), will increase the chance of accepting an alternative hypothesis when null hypothesis is in fact correct. True or False? Explain.

   False. Tails of t-distributions are wider, thus for a fixed significance level, critical value will be higher, and rejection region will be smaller, so chance of rejecting null (and accepting alternative) is decreased.

4. Some defendants in criminal proceedings plead guilty and are sentenced without a trial, whereas others plead innocent, are subsequently found guilty, and then are sentenced. In recent years legal scholars have speculated as to whether the sentences of those who plead guilty differ in severity from the sentences of those who plead innocent and subsequently judged guilty. Consider the accompanying data on a group of defendants from San Francisco County, all of whom were accused of robbery and had previous prison records ("Does it pay to plead guilty? Differential Sentencing and the functioning of criminal courts" Law and Society Rev. (1981-1982)).

	Plea	
	Guilty	Not Guilty
Number of judged guilty	n1=191	n2=64
Number sentenced to prison	101	56
Sample proportion	p_1'=.529	p_2'=.875

Solution:
Will be testing H0: p1 = p2 vs H1:p1 not equal to p2.
Can use normal approximation to Binomial since n1p1, n1(1-p1), n2p2, n2(1-p2) > 10.
Test statistics zobt=(p1'-p2')/sqrt(p1'(1-p1')/n1+ p2'(1-p2')/n2) =
(.529-.875)/sqrt(.529(1-.529)/191 + .875(1-.875)/64)=-.346/.055=6.29
Will reject H0 as zcrit=1.96.

5. The paper "Veterinary Health Care Market for Dogs" studied annual veterinary costs for households owning dogs. They recorded the annual veterinary costs (in $) for a random sample of 144 families. For these data:
$\Sigma_{i=1}^{144}(x_i) = 10{,}656$; $\Sigma_{i=1}^{144}(x_i)^2 = 1{,}017{,}344$
   a) Find the sample mean of the annual veterinary costs for one family owning a dog.
   b) Find the sample standard deviation of the annual veterinary costs of a family owning a dog.
   c) What will be distribution of the sample mean for a sample of size 144? With what parameters? Do we need to make any assumptions about distribution of X's?

   Solution:
   a) Sample mean = $\Sigma_{i=1}^{144}(x_i)/n$ = 10,656/144=74

   b) Sample st. deviation = sqrt{sum{xi- xbar}^2/n} =
   sqrt{sum(xi)^2/n-(sum(xi))^2/n^2}=
   sqrt{1,017,344/144-10,656^2/144^2}=39.86

   c) It will be approximately Normal (since sample size n=144 >30):

Normal(mu, 39.86/144), with mean mu equal to the true mean of the annual vet. costs for households owning dogs, and variance equal to square of sample st. deviation/144.

6. 9.07 had two graders this semester. The TA's job is to make sure the graders are equivalent.

a. On the midterm, the TA gave each grader half the exams to grade and then examined the results, shown below:

	Grader1	Grader2
N (number of exams graded)	20	22
Mean	59	63
Std. Dev.	10	12

Were the two graders significantly different at the alpha=.05 level?

Solution: use a two-sample hypothesis test
Sigma unknown, sample size less than 30 → use t-test.

t-obt = $\dfrac{63-59}{\sqrt{12^2/22+10^2/20}} = \dfrac{4}{\sqrt{6.55+5}} = \dfrac{4}{3.40} = 1.18$

df = min(19,21) = 19, t-crit = 2.093.

t-obt < t-crit, therefore, not significantly different.

b. On the final, the TA decided to have both graders grade the first 5 exams to make sure they were equivalent. Here are the scores for the 5 exams that each grader assigned:

	Grader1	Grader 2
Exam 1	50	47
Exam 2	67	62
Exam 3	57	59
Exam 4	71	68
Exam 5	40	43

Were the two graders significantly different at the alpha=.05 level?

Solution: use a matched pair test: Define $d_i = score_{i1} - score_{i2}$.
d = {3, 5, -2, 3, -3}.
Mean(d) = 1.2.

SD(d) = sqrt((3.24 + 14.44 + 10.24 + 7.84 + 17.64)/4) = 3.65
Use two-tailed t-test, df = 4, t-crit = 2.776.
t-obt = m / (SD/sqrt(N)) = 1.2 / (3.65 / sqrt(5)) = 0.74
t-obt < t-crit, therefore there is no significant difference between the graders.

c. Would the test be more powerful, less powerful, or the same if the TA compared the scores that the graders assigned to individual questions on the final? Explain your choice.

Solution: it would be more powerful, because there would be more samples to compare (a higher N). Also, the deviation in each score would probably be lower than in the exam overall.

7. Lamarck believed that physically changing an animal would affect its offspring. In one experiment, he cut the tail off rats and bred them to see if the children would be born with shorter tails. Each child rat has a mother and a father, so Lamarck tested 4 groups, each with 16 offspring. Here are the data:

	Length of Offspring's Tails
Normal Mother, Normal Father	m=45 $s^2 = 20$ n=16
Normal Mother, Tailless Father	m=38 $s^2 = 30$ n=16
Tailless Mother, Normal Father	m=47 $s^2 = 30$ n=16
Tailless Mother, Tailless Father	m=40 $s^2 = 20$ n=16

Complete the following table (and define A and B):

Solution: A is Father, B is Mother

SSA=784	dfA=1	MSA=784	Fobt,A=31.36
SSB=64	dfB=1	MSB=64	Fobt,B=2.56
SSAB=0	dfAB=1	MSAB=0	Fobt,AB=0
SSE=1500	dfE=60	MSE=25	
SST=2348	dfT=63		

a) Which factor(s) are significant (A, B, or both)?

Solution: A

b) Is there an interaction? Is it significant?

Solution: No interaction

c) Determine whether Tailless Mothers produce significantly different tail lengths in their offspring than Normal Mothers.

Solution:

1. using Fisher's post-hoc test: $t_{obt} = \frac{43.5 - 41.5}{\sqrt{25\left(\frac{1}{32} + \frac{1}{32}\right)}} = \frac{2}{1.25} = 1.6$

At df = 60, t-crit = 2.0, so this is not significantly different.

2. using Tukey's test: q2,60 = 2.83. HSD = 2.83*sqrt(25/32) = 2.50. This is greater than 43.5 – 41.5 = 2, so it is not significantly different.

d) What should Lamarck report about whether tailless mothers influence the offspring's tail length? Explain your reasoning.

Lamarck should conclude that there is no evidence at all for tailless mothers influencing the offspring's tail length. Lamarck set out to test whether cutting the tail would create shorter offspring tail lengths. Therefore, he should choose a one-tailed (no pun intended) test. Thus, any evidence for longer tail lengths should be considered to be due to chance.