Statistics 2000, Section 001, Midterm 2 (185 Points)

Friday, November 5, 2010

Part I: Text Answers

Your Name: _____

From: HW 5, Exercise 1, Q 2.111

Question 1: Two–Way Tables (**35 Points**)

Full–time and part–time college students. The U.S. Census Bureau provides estimates of numbers of people in the United States classified in various ways. Let's look at college students. The following table gives us data to examine the relation between age and full–time or part–time status. The numbers in the table are expressed as thousands of U.S. college students.

U.S. college students by age and status

	Status		
Age	Full-time	Part-time	Total
15–19	3388 20.67%	389 2.37%	3,777 23.05%
20–24	5238 31.96%	1164 7.10%	6,402 39.07%
25–34	1703 10.39%	1699 10.37%	3,402 20.76%
35 and over	762 4.65%	2045 12.48%	2,807 17.13%
Total	11,091 67.68%	5,297 32.32%	16,388 100.00%

Show your work!

1. (**5 Points**) What is the U.S. Census Bureau estimate of the *number* of full–time college students aged 15 to 19?

3,388,000 (5) (the numbers are in thousands!)

Note:
-3 if just 3,388

-1 if no work shown in 2., 3., 4.

2. **(10 Points)** Add the joint distribution of age and status to the table on the previous page. Show your calculations below.

First, calculate Totals: $3,388 + 5,238 + 1,703 + 762 = 11,091$ (in thousands)

...

$3,388 + 389 = 3,777$ (in thousands)

$11,091 + 5,297 = 16,388$ (in thousands)

(5)

Then, calculate percentages for joint distribution:

$\frac{3,388}{16,388} = 20.67\%$

(5)

3. **(10 Points)** What is the marginal distribution of age? Add these percentages to the table on the previous page. Show your calculations below. Also, display the results graphically via a bar chart.

Bar Chart: (4)

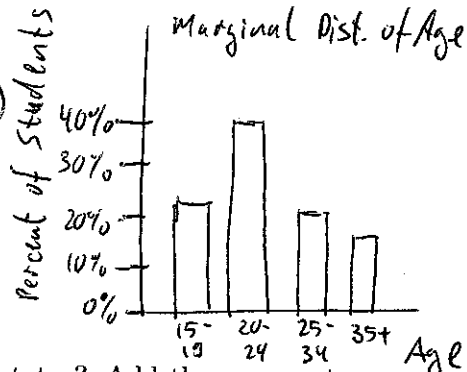
Marginal dist. of age: $\frac{3,777}{16,388} = 23.05\%$

$\frac{6,402}{16,388} = 39.07\%$

$\frac{3,402}{16,388} = 20.76\%$

$\frac{2,807}{16,388} = 17.13\%$

(6)



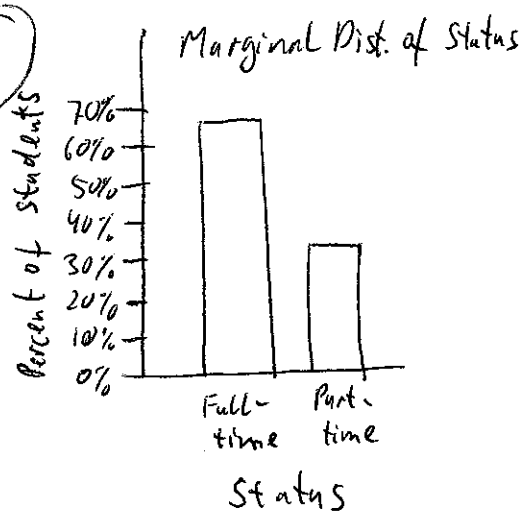
4. **(10 Points)** What is the marginal distribution of status? Add these percentages to the table on the previous page. Show your calculations below. Also, display the results graphically via a bar chart.

Bar Chart: (4)

Marginal dist. of status: $\frac{11,091}{16,388} = 67.68\%$

$\frac{5,297}{16,388} = 32.32\%$

(6)



-2 for each calculation error

+2 if correct formula (without any values)

Question 2: Probabilities, Means, and Variances of Random Variables (**50 Points**)

Based on past experience, an architect has determined the distribution for the number of times a drawing of a house must be examined by a client before it is accepted. Let the random variable X represent the number of times a drawing of a house must be examined by a client before it is accepted. The probability distribution of X is shown below:

Value of X	1	2	3	4	5
Probability	0.1	0.2	0.3	0.2	0.2

Show your work!

1. (**8 Points**) What is the average number (mean) μ_X of times a drawing of a house must be examined by a client before it is accepted?
   Answer: 3.2 times

$$\mu_X = \sum_{i=1}^{5} x_i p_i = 1 \cdot 0.1 + 2 \cdot 0.2 + 3 \cdot 0.3 + 4 \cdot 0.2 + 5 \cdot 0.2 = 3.2$$ (8)

2. (**8 Points**) Calculate the standard deviation σ_X of the random variable X.
   Answer: 1.25 times

$$\sigma_X^2 = \sum_{i=1}^{5} (x_i - \mu_X)^2 \cdot p_i = (1-3.2)^2 \cdot 0.1 + (2-3.2)^2 \cdot 0.2 + (3-3.2)^2 \cdot 0.3 + (4-3.2)^2 \cdot 0.2 + (5-3.2)^2 \cdot 0.2 = 1.56$$ (6)

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{1.56} = 1.25$$ (2)

3. (**8 Points**) What is the probability that a client examines a drawing of a house at least 3 times before it is accepted?
   Answer: 70 %

$$P(\text{at least 3 times}) = P(X \geq 3) = P(X=3) + P(X=4) + P(X=5) = 0.3 + 0.2 + 0.2 = 0.7 = 70\%$$ (8)

4. (**8 Points**) Mr. Klein is one of the clients of this architect. Mr. Klein has already inspected his drawing twice. What is the probability that he will accept the drawing upon the next examination, i.e., after exactly the third time?
   Answer: 42.86 %

Intuitive solution: Out of 100 drawings, 70 (which is 70%) will require 3 or more inspections. 30 of these will be accepted upon the next examination. So, the probability is $30/70 = 0.4286 = 42.86\%$ (8)

or: Statistical solution:

$$P(\text{accept on 3rd examination} \mid \text{at least 3 times}) = \frac{P(\text{accept on 3rd and at least 3 times})}{P(\text{at least 3 times})} = \frac{P(\text{accept on 3rd})}{P(\text{at least 3 times})} = \frac{P(X=3)}{P(X \geq 3)} = \frac{0.3}{0.7} = 0.4286 = 42.86\%$$ (8)

or:

Note: just 30%: -6

5. **(8 Points)** Let the random variable Y represent the number of times a drawing of a garage must be examined by a client before it is accepted. The architect determined that the average number (mean) μ_Y of times a drawing of a garage must be examined by a client before it is accepted is 1.8 times. Assuming a customer plans to get a house and a garage from the architect, what is the expected (average) number of times a customer will look at all drawings (of the house and the garage) before they are accepted?
Answer: 5.0 times

(8)

$\mu_{X+Y} = \mu_X + \mu_Y = 3.2 + 1.8 = 5.0$

6. **(10 Points)** Based on past data, the architect also determined that the standard deviation σ_Y of the random variable Y is 1.1 times. Moreover, some clients in general want to see many drawings of any building while other clients are satisfied after just a few drawings. So, no big surprise that the correlation between X and Y equals 0.90. What is the standard deviation for the number of times a customer will look at all drawings (of the house and the garage) before they are accepted?
Answer: 2.29 times

$\sigma^2_{X+Y} = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$

$= 1.25^2 + 1.1^2 + 2 \cdot 0.9 \cdot 1.25 \cdot 1.1$ (7)

$= 5.25$

$\sigma_{X+Y} = \sqrt{\sigma^2_{X+Y}} = \sqrt{5.25} = 2.29$ (3)

Grading: 0 incorrect result (no calculation, no P-notation)
+1 some P-notation
+2 some calculations
+2 some P-notation & some calculations
+4 correct result no calculation
+7 correct result + calculation
+8 correct result + calculation + P-notation

Question 3: Probability (**40 Points**)

For a road trip, a student places the following nine CDs into the glove compartment of his car:

- 5 *modern rock* CDs (Fallout Boy, Hawthorne Heights, The Used, Finger Eleven, Taking Back Sunday),
- 3 *pop* CDs (P!nk, Fergie, Gwen Stefani),
- 1 *American Idol* CD (Jordin Sparks).

Most direct solution(s) below; alternative solution on next page!

On his trip, the student blindly grabs a CD from the glove compartment, listens to it, and places it on the back seat when finished. Then he blindly grabs a second CD from the glove compartment. You should NOT comment on the musical taste of this student, but **answer each of the following questions separately. Show your work!** As a part of your answer, translate the everyday language into probability statements, using the proper notation, e.g., $P(A), P(A \text{ and } B), P(A \text{ or } B), P(A|B)$, etc., where A (and B) are the events of interest.

1. (**8 Points**) What is the chance that the FIRST CD will be a *pop* CD or the *American Idol* CD? The chance is 44.4 %

$P(\text{1st pop or 1st Idol}) = P(\text{1st pop}) + P(\text{1st Idol})$

$= \frac{3}{9} + \frac{1}{9} = \frac{4}{9} = 0.444 = 44.4\%$

2. (**8 Points**) What is the chance that the SECOND CD will be a *pop* CD or the *American Idol* CD? The chance is 44.4 %

$P(\text{2nd pop or 2nd Idol}) = P(\text{2nd pop}) + P(\text{2nd Idol})$

$= \frac{3}{9} + \frac{1}{9} = \frac{4}{9} = 0.444 = 44.4\%$

NOT: $\frac{1}{9} + \frac{1}{8} = \frac{17}{72}$

3. (**8 Points**) What is the chance that he will listen to Jordin Sparks as one of his two selections? The chance is 22.2 %

$P(\text{1st Jordin or 2nd Jordin}) = P(\text{1st Jordin}) + P(\text{2nd Jordin})$

$= \frac{1}{9} + \frac{1}{9} = \frac{2}{9} = 0.222 = 22.2\%$

or: ... $= P(\text{1st Jordin and 2nd other}) \text{ or } P(\text{1st other and 2nd Jordin})$

$= \frac{1}{9} \cdot \frac{8}{8} + \frac{8}{9} \cdot \frac{1}{8} = \frac{1}{9} + \frac{1}{9} = \frac{2}{9}$ (as before)

4. (**8 Points**) What is the chance that he will listen to **none** of the *pop* CDs? The chance is 41.7 %

$P(\text{1st not pop and 2nd not pop}) = P(\text{1st not pop}) \cdot P(\text{2nd not pop} \mid \text{1st not pop})$

$= \frac{6}{9} \cdot \frac{5}{8} = \frac{30}{72} = 0.417 = 41.7\%$

5. (**8 Points**) What is the chance that he will listen to **at least one** of the *modern rock* CDs? The chance is 83.3 %

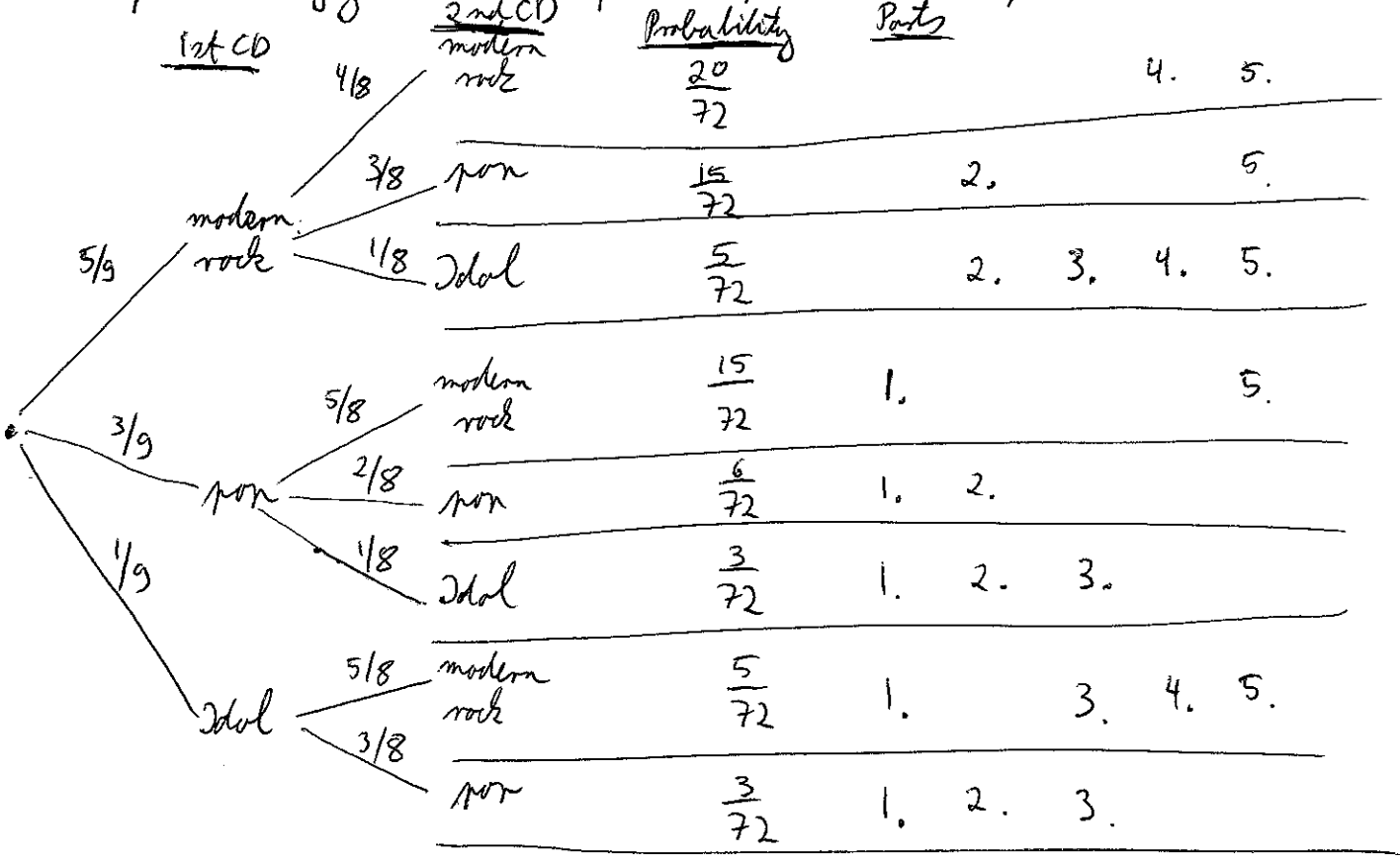
$P(\text{at least one modern rock}) = 1 - P(\text{both not modern rock})$

$= 1 - P(\text{1st not modern rock}) \cdot P(\text{2nd not modern rock} \mid \text{1st not modern rock})$

$= 1 - \frac{4}{9} \cdot \frac{3}{8} = \frac{72}{72} - \frac{12}{72} = \frac{60}{72} = 0.833 = 83.3\%$

Question 3: Alternative Solution

Construct a tree diagram, then read off the probabilities from the leaves and sum up accordingly to answer question parts 1.) to 5.):



1.) $P = \frac{15}{72} + \frac{6}{72} + \frac{3}{72} + \frac{5}{72} + \frac{3}{72} = \frac{32}{72} = \frac{4}{9} = 44.4\%$

2.) $P = \frac{15}{72} + \frac{5}{72} + \frac{6}{72} + \frac{3}{72} + \frac{3}{72} = \frac{32}{72} = \frac{4}{9} = 44.4\%$

3.) $P = \frac{5}{72} + \frac{3}{72} + \frac{5}{72} + \frac{3}{72} = \frac{16}{72} = \frac{2}{9} = 22.2\%$

4.) $P = \frac{20}{72} + \frac{5}{72} + \frac{5}{72} = \frac{30}{72} = 41.7\%$

5.) $P = \frac{20}{72} + \frac{15}{72} + \frac{5}{72} + \frac{15}{72} + \frac{5}{72} = \frac{60}{72} = 83.3\%$

Statistics 2000, Section 001, Midterm 2 (185 Points)

Friday, November 5, 2010

Part II: Multiple Choice Questions

Your Name: _____

Question 4: Multiple Choice Questions **(60 Points)**

Mark your answer for each multiple choice question in the table below. There is only one correct answer for each question. Each correct answer is worth 4 points.

Question	(a)	(b)	(c)	(d)	Question	(a)	(b)	(c)	(d)
1	○	○	●	○	11	●	○	○	○
2	○	●	○	○	12	●	○	○	○
3	○	○	○	●	13	○	○	●	○
4	○	○	○	●	14	○	○	●	○
5	○	○	●	○	15	○	○	●	○
6	○	●	○	○					
7	○	●	○	○					
8	○	○	●	○					
9	○	●	○	○					
10	●	○	○	○					

based on:

- 5 questions based on "Online Quizzes" for Chapters 2, 3 & 4
- 5 questions based on previous versions of Stat 2000, Midterm 2 (Fall 2009 & Spring 2010)

Stat 2000, Midterm 2, Question 4 — Solutions

1. (c) This is a double–blind experiment. Neither the patients nor the physician know which patient gets which drug. Only the assistant has a list with the information who gets which drug.

2. (b) Herbal tea is the explanatory variable. Obviously, there are many confounding variables here, but that wasn't the question.

3. (d) 3 types of glue $\times$ 2 bindings = 6 combinations (or treatments) overall.

4. (d) The participants were guaranteed confidentiality and anonymity — so neither of these can be compromised. If a researcher wants to follow up, even only if to see who has replied, the researcher cannot guarantee anonymity before the study.

5. (c) We have to use digit 1 (for Becker), ignore the next 1 (as we want a SRS of size 3), then use digit 7 (for Taylor), and then use digit 9 (for Weiss).

6. (b) Combining Rules 1 and 2 for means says $\mu_{X-Y} = \mu_X - \mu_Y$. So the mean difference (male − female) is $7.25 - 6.50 = 0.75$.

7. (b) Since genders of children from different pregnancies are independent, we can multiply the probability of a boy on any pregnancy by itself, so we have $0.513 \cdot 0.513 = 0.263$, or about 26.3%.

8. (c) There are 11 plums among the total 44 fruit, and since we are picking at random, each of the fruits in the refrigerator is equally likely to be chosen, so the probability of a plum is 11/44 or 1/4.

9. (b) By the addition rule, since Freshman and Sophomore are disjoint, P (Freshman or Sophomore) = P (Freshman) + P (Sophomore) $= 0.35 + 0.30 = 0.65$. The event Junior or Senior is the complement of the event Freshman or Sophomore, so its probability is $1 - 0.65 = 0.35$ (or 35%) by the rule for complements.

10. (a) This is the definition of randomness. Text reference: MMC, page 239.

11. (a) mean $= 0 \cdot 0.6 + 1 \cdot 0.1 + 2 \cdot 0.1 + 3 \cdot 0.1 + 4 \cdot 0.1 = 10 \cdot 0.1 = 1.0$

12. (a) P = area from truncated triangle from 0 to 1 = 1− area of triangle from 1 to 2 $= 1 - 1/2 \cdot$ width $\cdot$ height $= 1 - 1/2 \cdot 1 \cdot 1/2 = 1 - 1/4 = 3/4 = 0.75$.

13. (c) We observe pairs of outcomes — and there are four different pairs.

14. Just (c) Since the tosses of the penny and nickel do not influence each other, any event concerning the outcome for the penny is independent of any event concerning the outcome for the nickel. So the event A, the penny shows a head, is independent of the event B, the nickel shows a tail.

15. (c) The explanatory variables in an experiment are often called factors. Many experiments study the joint effects of several factors. This experiment is studying the effect of two factors, type of drug and dosage of drug on the response, which is concentration in the blood. Each treatment is formed by combining the specific values (levels) of each of the factors. Since there are two drugs and three dosages for each, there are $3 \cdot 2 = 6$ treatments in this experiment, which is why the subjects were divided into six groups, one for each possible treatment.