Statistics 1040, Section 004, Midterm 2 (200 Points)

Friday, March 26, 2004

Your Name: _____

from: Stat 1040, Test 1, June 26, 2000 (Summer 2000), Question 6

Question 1: Regression (**50 Points**)

-2 each calculation error

From the subjects (all men) in a health survey, the following data were collected:

X Average height = 68 inches SD = 2.5 inches

Y Average blood pressure = 120 mm SD = 15 mm

Correlation = -0.2.

The scatter diagram is football–shaped.

Show your work!

-4 if x, y flipped

-2 if x, y not specified

1. (**15 Points**) Find the regression equation for predicting blood pressure from height.

$$\text{slope} = r \cdot \frac{SD_y}{SD_x} = -0.2 \cdot \frac{15}{2.5} = -1.2 \quad (6)$$

$$\text{intercept} = \text{avg}\,y - \text{slope} \cdot \text{avg}\,x = 120 - (-1.2) \cdot 68 = 120 + 81.6 = 201.6 \quad (6)$$

equation: $\boxed{\text{blood pressure} = 201.6 - 1.2\ \text{height}}$ or $\boxed{y = 201.6 - 1.2 \cdot x}$ (3)

2. (**10 Points**) Using your regression equation, estimate the blood pressure of a man who is 73 inches tall.

-2 for old method, correct result

-8 for old method, incorrect result

predicted blood pressure of 73" tall men $= 201.6 - 1.2 \cdot 73$

$= 201.6 - 87.6 = 114$ mm

3. (**15 Points**) Find the r.m.s. error for predicting the blood pressure from the height.

-5 for each major mistake, e.g. SD_x instead of SD_y, $\sqrt{}$ of everything, r instead of r^2, etc.

$$\text{r.m.s. error} = \sqrt{1-r^2} \cdot SD_y$$

$$= \sqrt{1-(-0.2)^2} \cdot 15$$

$$= \sqrt{1-0.04} \cdot 15 = \sqrt{0.96} \cdot 15 = 14.69\ \text{mm} \approx 14.7\ \text{mm}$$

4. (**10 Points**) The correlation coefficient tells us that, on average, taller men have **higher** / **lower** blood pressures than shorter men, and that the relationship between blood pressure and height is quite **strong** / **weak**. (**Circle the correct word in each pair of choices.**) (5) (5)

from: FPP, Chapter 17, p. 305, Question 7

and: Stat 1040, Midterm 2, Nov. 8, 2002, Question 1

Question 2: The Expected Value and the Standard Error (**40 Points**)

A hundred draws are made at random with replacement from the box

1	2	3	4	5	6

.

Show your work!

-2 each calculation error

1. (**10 Points**) If the sum of the draws is 321, what is their average?
   It is: 3.21

$$\text{avg} = \frac{\text{sum of draws}}{\text{\# draws}} = \frac{321}{100} = 3.21 \quad (10)$$

2. (**10 Points**) If the average of the draws is 3.78, what is the sum?
   It is: 378

$$\text{sum of draws} = \text{avg} \cdot \text{\# draws} = 3.78 \cdot 100 = 378 \quad (10)$$

3. (**20 Points**) Using the normal curve, estimate the chance that the average of the draws is between 3 and 4.
   It is: 99.7%

Workbook: "The average will be between 3 and 4 if the sum is between 300 and 400. The average of the box is 3.5, and the SD is 1.7, so the expected value of the sum is 350 and the standard error is 17. The area is about 99.7%."

Additional details:

avg of 3 → sum of 300 (2)

avg of 4 → sum of 400 (2)

$$\text{box avg} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5 \quad (2)$$

$$\text{box SD} = \sqrt{\frac{(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2}{6}}$$

$$= \sqrt{\frac{17.5}{6}} = \sqrt{2.917} = 1.7 \quad (2)$$

$$EV_{sum} = 100 \cdot 3.5 = 350 \quad (2)$$

$$SE_{sum} = \sqrt{100} \cdot 1.7 = 17 \quad (2)$$

s.u.: $\frac{300-350}{17} = -2.94$ (2)

$\frac{400-350}{17} = 2.94$ (2)

area between −2.95 to 2.95:

99.68% ≈ 99.7% (4)

from: Stat 1040, Midterm 2, March 27, 2002, Question 1

Question 3: Chances and Probabilities (**40 Points**)

I have a bag with 20 balls in it: 10 are red, 8 are blue, and 2 are green.
Show your work!

1. (**10 Points**) If I draw one ball at random from the bag, what is the chance that I get a red ball or a green ball?
   It is: 60 %

-2 each calculation error

chance of red: $\frac{10}{20} = \frac{1}{2} = 0.5$ (3)

chance of green: $\frac{2}{20} = \frac{1}{10} = 0.1$ (3) → mutually exclusive → addition rule (4)

chance of red or green = $\frac{10}{20} + \frac{2}{20} = \frac{12}{20} = \frac{3}{5} = 0.6 = 60\%$

2. (**10 Points**) If I draw two balls at random **without replacement**, what is the chance that I get a red ball, followed by a green ball?
   It is: 5.3 %

chance first red: $\frac{10}{20}$ (3)

chance second green, given first red: $\frac{2}{19}$ (3) → dependent

chance first red and second green = $\frac{10}{20} \cdot \frac{2}{19} = \frac{1}{19} = 0.0526 \approx 5.3\%$ (multiplication rule (4))

3. (**10 Points**) If I draw three balls at random **with replacement**, what is the chance that I get all three red balls?
   It is: 12.5 %

chance first red: $\frac{10}{20} = \frac{1}{2}$ (2)

chance second red: $\frac{10}{20} = \frac{1}{2}$ (2) → independent

chance third red: $\frac{10}{20} = \frac{1}{2}$ (2)

chance all three red = $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

$= 0.125 = 12.5\%$

multiplication rule (4)

4. (**10 Points**) If I draw three balls at random **with replacement**, what is the chance that I get at least one red ball?
   It is: 87.5 %

chance first not red: $\frac{10}{20} = \frac{1}{2}$ (2)

chance second not red: $\frac{10}{20} = \frac{1}{2}$ (2) → independent

chance third not red: $\frac{10}{20} = \frac{1}{2}$ (2)

multiplication rule (2)

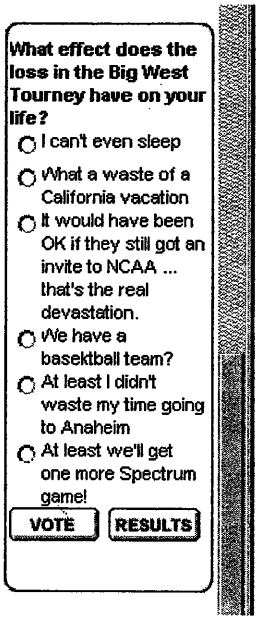
chance all three not red = $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

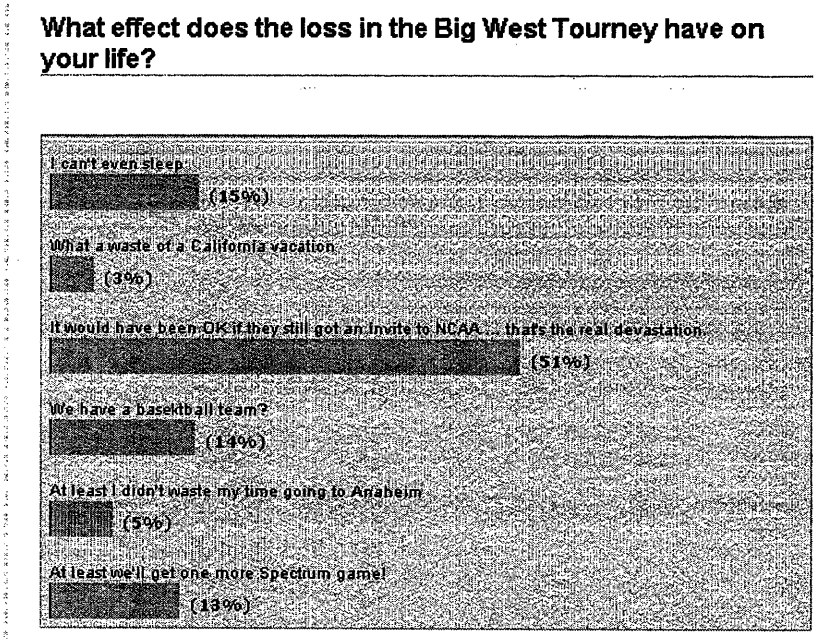
chance at least one red = $1 - \frac{1}{8} = \frac{7}{8} = 0.875 = 87.5\%$

↑ opposites rule (2)

Question 4: Sampling (**40 Points**)

What effect does the loss in the Big West Tourney have on your life? This was the question asked in a recent survey posted on the Utah Statesman Online Web page. A snapshot of the question and the survey results are shown below:





1. (**20 Points**) Whereas more than 50% of the students think that not being invited to the NCAA was the real devastation, 14% seem to be surprised to hear that USU has a basketball team. Interesting... But can we really take the results of this survey seriously? **Yes – seriously** / **no – not seriously.** Circle your answer and explain why or why not!

(10)

many reasons: (10)

- voluntary answers: only students that knew about this Web-site could access it and had a chance to indicate their opinion
- in general, students with a strong opinion are more likely to vote at all
- students could vote more than once!

etc...

2. (**20 Points**) Explain how **you** would organize a survey to obtain students' opinion on the effect of the loss in the Big West Tourney on their life.

(10)

Conduct a simple random sample (SRS): get names / social security numbers from all students, let a computer draw a few hundred of these names / SSNs, then interview those students (try everything to get their answers: write them, call them, e-mail them! – and perhaps offer a prize for responding!)

(10) for explanation

from: FPP, Chapter 20, p. 372, Question 7

Question 5: Chance Errors in Sampling **(30 Points)**

Five hundred draws are made at random from the box

60,000 × 0	20,000 × 1

True or false, and explain:

1. **(5 Points)** True / false: The expected value for the percentage of 1's among the draws is exactly 25%.

2. **(5 Points)** True / false: The expected value for the percentage of 1's among the draws is around 25%, give or take 2% or so.

3. **(5 Points)** True / false: The percentage of 1's among the draws will be around 25%, give or take 2% or so.

4. **(5 Points)** True / false: The percentage of 1's among the draws will be exactly 25%.

5. **(5 Points)** True / false: The percentage of 1's in the box is exactly 25%.

6. **(5 Points)** True / false: The percentage of 1's in the box is around 25%, give or take 2% or so.

calculation:

$\text{box avg} = \frac{20,000}{80,000} = \frac{1}{4} = 0.25 (= 25\%)$

$\text{box SD} = \sqrt{\frac{20,000}{80,000} \cdot \frac{60,000}{80,000}} = \sqrt{\frac{1}{4} \cdot \frac{3}{4}}$

$= \sqrt{\frac{3}{16}} = 0.433$

$EV_{sum} = 500 \cdot \frac{1}{4} = 125$

$SE_{sum} = \sqrt{500} \cdot 0.433 = 9.68$

$EV_{\%} = 25\%$

$SE_{\%} = \frac{9.68}{500} \cdot 100\% = 1.94\% \approx 2\%$

(3) 1. see calculation: (2)
($EV_{\%} = 25\%$)

(3) 2. we know exactly the expected value for the percentage of 1's among the draws (which is 25% - no give or take) (2)

(3) 3. see calculation: (2)
(close to $EV_{\%} = 25\%$, but give or take of about $SE_{\%} = 2\%$)

(3) 4. the percentage of 1's most likely will not be exactly 25% (2)
(but it will be relatively close to 25%)

(3) 5. see calculation: (2)
(box avg = fraction of 1's = 25%)

(3) 6. we know exactly the percentage of 1's in the population (i.e., box)
(which is 25% - no give or take) (2)