

Statistics 1040, Section 004 Midterm 2 (200 Points)

Friday, April 1, 2005

Your Name: _____

Question 1: Sampling (30 Points)

The pictures below show the The Utah Statement Online Poll regarding ASUSU elections as of 03/27/2005:

Will you be voting in the ASUSU primary or final elections?

Yes, it's important for the future of this university.

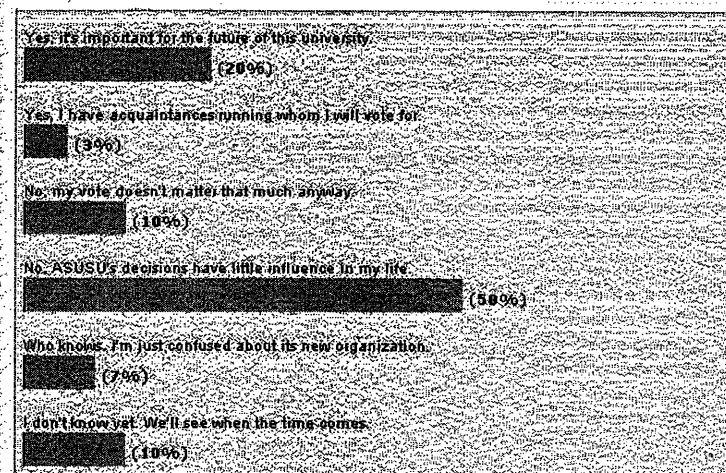
Yes, I have acquaintances running whom I will vote for.

No, my vote doesn't matter that much anyway.

No, ASUSU's decisions have little influence in my life.

Who knows, I'm just confused about its new organization.

I don't know yet, we'll see when the time comes.



Based on this poll, 50% of the USU students will not participate in the upcoming elections because ASUSU's decisions have little influence in their lives. Do you think the results of this poll objectively represent the opinion of ALL USU students? Answer yes or **no** and give **three arguments** to justify your answer, based on the facts given, *not* on your personal opinion.

9

- 1) Students who feel strongly are more likely to vote + these students may vote multiple times on this web poll.
 - 2) The source will create a bias - what type of students reads this web page (instead of the paper copy of the Statesman).
 - 3) Students must have access to a computer (Web browser at the time this question was posted (those on a field trip, excursion, or conference would not be able to vote).
 - 4) Who says that everyone who reads this web page (& votes) is a USU student??
- + 7 for each reasonable argument
- 3 for repeating the same in different words
- 5 if not a real argument

Question 2: Probability / EV, ES, and Normal Curve (90 Points)

Professor J.S. recently experimented with homegrown chili plants. From 4 seeds that were planted, one seed germinated. Use this chance (25%) as the basis for all calculations in this question and assume that seeds germinate independently from each other.

Show your work!

-2 for each calculation error
(or no final result in %)
-10 if % > 100% or % < 0%

Part 1 (40 Points)

In a second experiment, J.S. planted 4 more seeds. Determine the following chances:

1. (15 Points) The chance that all 4 of these 4 seeds will germinate is about 0.39 %.

first germinates: $\frac{1}{4}$
 second germinates: $\frac{1}{4}$
 third germinates: $\frac{1}{4}$
 fourth germinates: $\frac{1}{4}$
 all 4 germinate: $\frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \left(\frac{1}{4}\right)^4 = 0.0039 = \underline{\underline{0.39\%}}$

independent (8)
 multiplication rule (7)

-6 if not using $\frac{1}{4}$ in this part

2. (15 Points) The chance that none of these 4 seeds will germinate is about 31.64 %.

first does not germinate: $\frac{3}{4}$
 second does not germinate: $\frac{3}{4}$
 third does not germinate: $\frac{3}{4}$
 fourth does not germinate: $\frac{3}{4}$
 all 4 do not germinate: $\frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \left(\frac{3}{4}\right)^4 = 0.3164 = \underline{\underline{31.64\%}}$

independent (8)
 multiplication rule (7)

-6 if using $\frac{1}{4}$ in this part

3. (10 Points) The chance that at least 1 of these 4 seeds will germinate is about 68.36 %.

"at least 1" is opposite of "none":

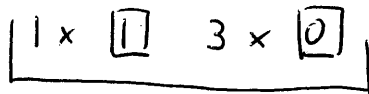
at least 1 germinates: $1 - \left(\frac{3}{4}\right)^4 = 1 - 0.3164 = 0.6836 = \underline{\underline{68.36\%}}$

opposite rule (9)

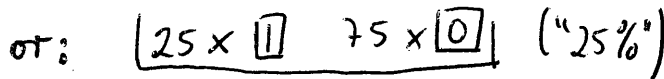
Part 2 (50 Points)

To meet his family's annual need in chilis, J.S. plans to plant 100 more seeds later in the Spring.

1. (10 Points) Indicate the box model.



draws: 100



("25%")

1: seed germinates

0: seed does not germinate

-2 for each calculation error

-3 for slightly incorrect number of [0] / [1]'s in box

-5 if box given as [0] [1] etc.

-7 if box contains something else than [0] / [1]'s

-2 if # draws missing or incorrect

2. (20 Points) The expected number of these 100 seeds that will germinate is 25 with a standard error of about 4.33.

$$\text{box avg} = \frac{1}{4} = 0.25$$

$$\text{box SD} = \sqrt{\frac{1}{4} \cdot \frac{3}{4}} = \sqrt{\frac{3}{16}} = \sqrt{0.1875} = 0.433$$

$$EV_{\text{sum}} = 100 \cdot 0.25 = \underline{\underline{25}}$$

$$SE_{\text{sum}} = \sqrt{100} \cdot 0.433 = 10 \cdot 0.433 = \underline{\underline{4.33}}$$

-2 for each minor mistake

-5 for each major mistake (or step missing)

3. (20 Points) To meet his family's annual need in chilis, 20 chilis that germinate would be sufficient. The chance that at least 20 of these 100 seeds will germinate is about 87.5 %.

$$\text{s.u.} = \frac{20 - 25}{4.33} = -1.15$$



20 25
-1.15 0 s.u.

area between -1.15 and 1.15: $74.99\% \approx 75\%$

area above -1.15: $\frac{75\%}{2} + 50\% = 37.5\% + 50\% = \underline{\underline{87.5\%}}$

-4 for each incorrect curve parameter, i.e., anything else than EV & SE

-4 for incorrect s.u.

-4 for incorrect table value (not 74.99%)

-4 for incorrect area under the curve (not 87.5%)

from: Stat 1040, Fall 2001, Final Test, December 13, 2001, Question 2 b & c

Question 3: Regression (40 Points)

A selection of 65 varieties of cereal were tested for calories and sodium (in milligrams) for an one-cup serving. The results can be summarized as follows:

- Y Average sodium = 240 mg; SD = 131 mg;
X Average calories = 149 calories; SD = 62 calories; $r = 0.53$.

-2 each calculation error

Show your work!

-2 if x, y flipped

-2 if x, y not specified

1. (10 Points) Find the equation of the regression line for predicting number of mg sodium in an one-cup serving of cereals from calories.

$$\text{slope} = r \cdot \frac{SD_y}{SD_x} = 0.53 \cdot \frac{131}{62} = 1.12 \quad (4)$$

$$\text{intercept} = \text{avg } y - \text{slope} \cdot \text{avg } x = 240 - 1.12 \cdot 149 = 240 - 166.88 = 73.12 \quad (4)$$

$$\text{regression equation: sodium} = 73.12 + 1.12 \cdot \text{calories} \quad (2)$$

$$\text{or } y = 73.12 + 1.12 \cdot x$$

2. (10 Points) Predict the number of mg sodium in an one-cup serving of cereals that has 200 calories per cup.

-2 for old method, correct result
-8 for old method, incorrect result

$$\text{for 200 calories: sodium} = 73.12 + 1.12 \cdot 200 = 73.12 + 224 = \underline{\underline{297.12}}$$

3. (10 Points) Find the r.m.s. error for predicting mg sodium from calories.

$$\begin{aligned} \text{r.m.s. error} &= \sqrt{1-r^2} \cdot SD_y &= \sqrt{0.7191} \cdot 131 \\ &= \sqrt{1-0.53^2} \cdot 131 &= 0.848 \cdot 131 \\ &= \sqrt{1-0.2809} \cdot 131 &= \underline{\underline{111.1}} \end{aligned}$$

-4 for each major mistake,
e.g. SD_x instead of SD_y ,
 $\sqrt{\text{of everything}}$, r instead
of r^2 etc.

4. (10 Points) Explain why it would not be a good idea to use the information in the question to estimate the amount of sodium for a cereal with 350 calories per cup.

$$350 \text{ calories: } \frac{350-149}{62} = 3.2 \text{ s.u.}$$

350 is more than 3 s.u. above the average (of 149 calories);

this is extrapolation and the result will be meaningless

(perhaps this cereal contains 4 big chocolate chips that are high in calories, but may be low on sodium)

(5)

from: FPP, p. 328, Chapter 18, Review Exercise 8

Question 4: Normal Approximation for Probability Histograms (40 Points)

A coin is tossed 100 times. True or false, and briefly explain.

Answer and explain each of the following questions separately!

H: 1

T: 0

1. (10 Points) The expected value for the number of heads is 50.

7 True False

Explanation:

$$\text{box avg} = \frac{1}{2}$$

$$\text{box SD} = \sqrt{\frac{1}{2} \cdot \frac{1}{2}} = \frac{1}{2}$$

$$EV_{\text{sum}} = 100 \cdot \frac{1}{2} = 50$$

$$SE_{\text{sum}} = \sqrt{100} \cdot \frac{1}{2} = 5$$

3 correct, $EV_{\text{sum}} = 50$

box:

0	1
---	---

draws: 100

-5 if no box model /
EV/SE calculations

2. (10 Points) The expected value for the number of heads is 50, give or take 5 or so.

True False 7

Explanation:

False, the EV_{sum} is exactly 50, no give or take here 3

3. (10 Points) The number of heads will be 50.

True False 7

Explanation:

False, the number of heads most likely will not be exactly 50, but it will be relatively close to 50 3

4. (10 Points) The number of heads will be around 50, give or take 5 or so.

7 True False

Explanation:

Correct, as calculated in 1.) $EV_{\text{sum}} = 50$ and $SE_{\text{sum}} = 5$
(and we usually use the SE_{sum} for the "give or take" part) 3