

MATH 181 FIRST EXAM PRACTICE

SPRING 2019

Name: ANSWER KEY

- Write your **full name** on the line above.
- Show your work. Incorrect answers with work can receive partial credit.
- Attempt every question; showing you understand the question earns some credit.
- If you run out of room for an answer, continue on the back of the page. Before doing so, write “see back” with a circle around it.
- You can use 1 page (front and back) of notes.
- You can use (and probably need) a calculator.
- You can use the Geogebra Scientific Calculator instead of a calculator. You need to put your phone on **airplane mode** and then within the application, start **exam mode**; you should see a green bar with a timer counting up.
- If a question is confusing or ambiguous, please ask for clarification; however, you will not be told how to answer the question.
- **Box your final answer.**
- A formula sheet is attached to this test.

Do not write in this grade table.

Question:	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Total
Points:	10	10	10	10	10	10	10	10	80
Score:									

Sample statistics: n = sample size x_i = the i th value in a sample \bar{x} = sample mean s = sample standard deviation

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

 Q_1 = first quartile m = median Q_3 = third quartileIQR = inter-quartile range = $Q_3 - Q_1$

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Population parameters: μ = population mean σ = population standard deviation**Probability:** Ω = set of all possible equally likely outcomes A = event A , a set of outcomes A^c = The complement of A B = event B , another set of outcomes $|A|$ = size of set, number of outcomes in A $P(A)$ = probability of A $P(A \text{ AND } B)$ = probability of both A and B $P(A \text{ OR } B)$ = probability of either A or B (or both) $P(A|B)$ = probability of A given B

$$P(A) = \frac{|A|}{|\Omega|}$$

$$0 \leq P(A) \leq 1$$

$$P(A \text{ AND } B) = P(A) \cdot P(B|A)$$

$$P(A \text{ OR } B) = P(A) + P(B) - P(A \text{ AND } B)$$

$$P(A^c) = 1 - P(A)$$

$$A, B \text{ are disjoint (mutually exclusive)} \iff P(A \text{ AND } B) = 0$$

$$A, B \text{ are non-disjoint} \iff P(A \text{ AND } B) > 0$$

$$A, B \text{ are exhaustive} \iff P(A \text{ OR } B) = 1$$

$$A, B \text{ are complements} \iff A, B \text{ are disjoint and exhaustive} \iff B = A^c$$

$$A, B \text{ are independent} \iff P(A \text{ AND } B) = P(A) \times P(B) \iff P(A|B) = P(A)$$

Random variables and distributions: X = random variable x_i = the i th possible value of X . (Notice different meaning here vs. sample statistics.) k = number of possible values of X . $E(X) = \mu$ = expected value of X σ = standard deviation of X

$$\mu = \sum_{i=1}^k x_i \cdot P(X = x_i)$$

$$\sigma = \sqrt{\sum_{i=1}^k (x_i - \mu)^2 \cdot P(X = x_i)}$$

Q1. (10 points) An urn contains marbles. Each marble has a color and a pattern. The frequencies are shown in the contingency table.

		R	G	B	
		red	green	blue	total
D	dotted	89	31	22	142
S	striped	22	41	97	160
C	checkered	16	36	42	92
	total	127	106	161	394

(a) What is the probability that a random marble is green?

$$P(G) = \frac{106}{394} \approx 0.269$$

(b) What is the probability that a random marble is striped and green?

$$P(S \text{ and } G) = \frac{41}{394} \approx 0.104$$

(c) What is the probability that a random marble is striped or green?

$$P(S \text{ or } G) = \frac{22+41+97+31+36}{394} = \frac{160+106-41}{394} \approx 0.571$$

(d) What is the probability that a random marble is striped given it is green?

$$P(S|G) = \frac{41}{106} \approx 0.387$$

(e) What is the probability that a random marble is green given it is striped?

$$P(G|S) = \frac{41}{160} \approx 0.256$$

- Q2.** (10 points) Amira ran a study to see whether a rooting hormone increased the survival rate of roma-tomato cuttings (some plants can propagate by slicing off a branch and planting it). After collecting 18 cuttings, she dipped 9 (randomly selected) cuttings in rooting hormone and the others in water. After 4 weeks, she noted how many from each group survived.

	survived	died
hormone	6	3
water	5	4

- (a) Is this study observational or experimental?

Experimental

- (b) What is the sample?

18 cuttings

- (c) What are the cases (individuals)?

the cuttings

- (d) What are the variables?

survival and dipping substance

- (e) Could a causal relationship be established by the study design?

yes

- (f) Was blinding used in this study?

no

- (g) Which treatment (hormone or water) had a higher proportion surviving after 4 weeks?

The hormone treatment had a higher proportion. $\frac{6}{9} > \frac{5}{9}$

- (h) Describe the null hypothesis in the context of this study.

H_0 : The difference is just due to chance.

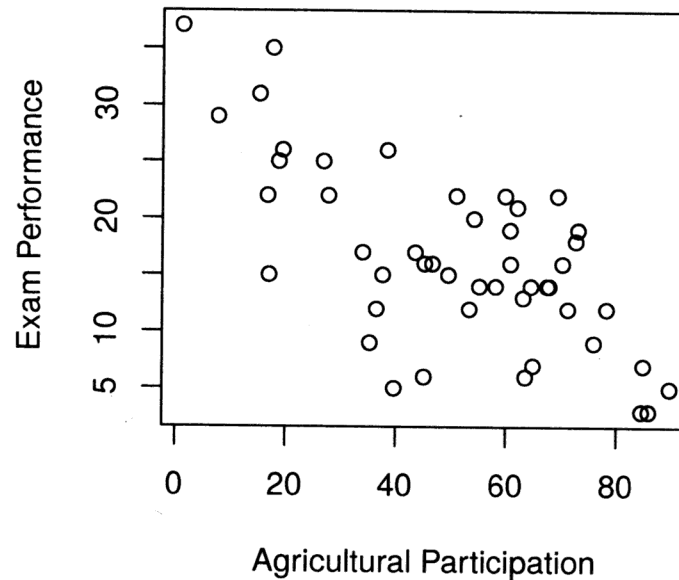
- (i) Describe the alternative hypothesis in the context of this study.

H_A : Hormones cause a different survival rate than water.

- (j) Has Amira shown, beyond a reasonable doubt, that hormone is more effective than water? Why or why not?

No. It seems like chance could easily account for the difference.

- Q3.** (10 points) In 1888, the Swiss government collected data on 47 French-speaking provinces. Two of the variables were the percent of men engaged in agriculture and the percent of draftees earning top marks on an examination.



- (a) Which variable is the implied explanatory variable?

Agricultural participation

- (b) Which variable is the implied response variable?

Exam performance

- (c) What kind of association (positive, negative, or none) exists between these two variables?

Negative

- (d) A friend suggests this study shows that agricultural participation causes lower exam performance. Is this a reasonable conclusion? Why or why not?

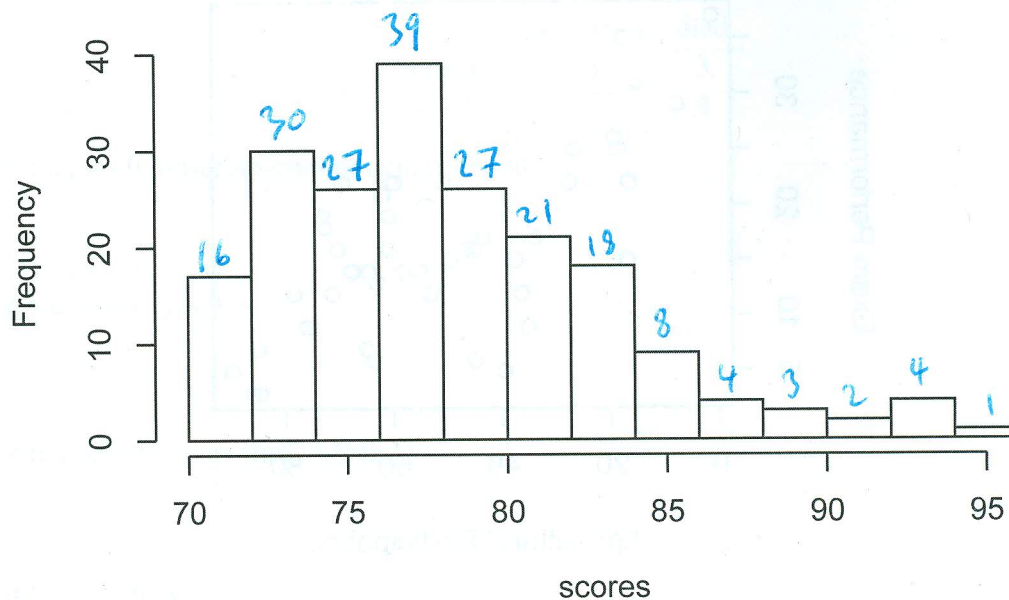
This is not reasonable. Observational studies do not show causal relationships.

- (e) What is a possible confounding variable?

Education, wealth, socioeconomic status, cultural values...

- Q4.** (10 points) A teacher has given an exam to 200 students, whose scores are shown in the histogram below. Please assume that the scores are almost all non-integer (this is meant to simplify the analysis by removing inclusive/exclusive concerns).

Histogram of scores



- (a) Estimate the proportion of students who scored between 70 and 74.

$$\approx \frac{46}{200}$$

$$\boxed{23\%}$$

- (b) Estimate the median.

$$\approx 77$$

- (c) Is the mean higher or lower than the median? Why?

The mean is higher than the median; the distribution is skewed right.

- (d) Estimate the first quartile, Q_1 .

$$\approx 74.5$$

- (e) Estimate the third quartile, Q_3 .

$$\approx 81$$

- (f) Which of the following choices is the best estimate of the standard deviation.

- ☐ 1
☒ 5
☐ 30
☐ 80
☐ 95

Q5. (10 points) Gary measured the masses (in grams) of 8 random oranges.

259 254 259 267 244 263 242 300

(a) Find the sample mean (\bar{x}).

$$\frac{259 + 254 + 259 + 267 + 244 + 263 + 242 + 300}{8} = 261$$

$$\bar{x} = 261$$

(b) Find the sample standard deviation (s).

x	$(x - \bar{x})^2$
259	4
254	49
259	4
267	36
244	289
263	4
242	361
300	1521

$$\sum (x - \bar{x})^2 = 2268$$

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{2268}{7}} = 18$$

$$s = 18$$

(c) Make a box plot. Label the features (with numbers).

sort

242
244
254
259
259
263
267
300

$Q_1 = 249$

median = 259

$Q_3 = 265$

$$IQR = Q_3 - Q_1$$

$$IQR = 265 - 249$$

$$IQR = 16$$

$$Q_1 - 1.5 \times IQR$$

$$249 - 24$$

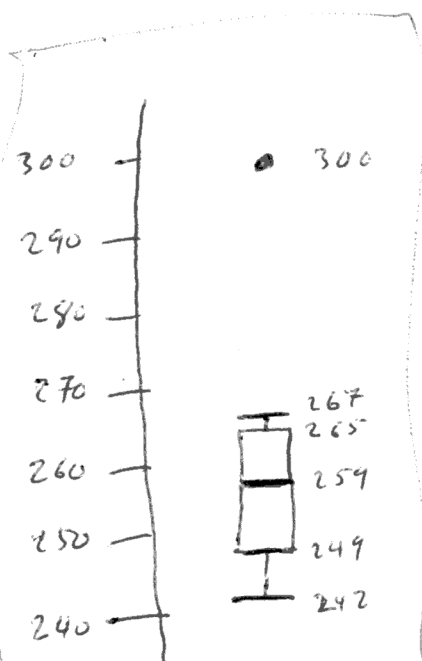
$$225$$

$$Q_3 + 1.5 \times IQR$$

$$265 + 24$$

$$289$$

to check for outliers



Q6. (10 points) Jenn has a weighted die that rolls with probability distribution below. Let random variable X represent the result of a roll.

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
1	0.10	0.1
2	0.15	0.3
3	0.15	0.45
4	0.15	0.6
5	0.15	0.75
6	0.30	1.8

$$\sum_{i=1}^6 x_i \cdot P(X = x_i) = 4$$

$$\mu = 4$$

$(x_i - \mu)^2 \cdot P(X = x_i)$
0.9
0.6
0.15
0
0.15
1.2

$$\sigma^2 = 3$$

$$\sigma \approx 1.73$$

(a) What is $P(X = 6)$?

$$0.30$$

(b) Evaluate $P(2 \leq X \leq 5)$.

$$0.15 + 0.15 + 0.15 + 0.15 = 0.6$$

(c) Evaluate the mean of the probability distribution.

$$\mu = 4 = 0.1 + 0.3 + 0.45 + 0.6 + 0.75 + 1.8$$

(d) Evaluate the standard deviation of the probability distribution.

$$\sigma \approx 1.73$$

(e) Assume multiple rolls are independent, where X_i is the result of the i th roll. Evaluate the probability $P(X_1 = 6 \text{ AND } X_2 = 6)$. In other words, what is the chance of rollings two 6s in a row?

$$(0.3)(0.3) = 0.09$$

(f) Evaluate $P(X_1 \neq 6 \text{ AND } X_2 \neq 6 \text{ AND } X_3 \neq 6)$. In other words, what is the chance of rolling thrice and getting no 6s?

$$0.7 \times 0.7 \times 0.7 = 0.343$$

(g) Evaluate $P(X_1 = 6 \text{ OR } X_2 = 6 \text{ OR } X_3 = 6)$. In other words, what is the chance of rolling thrice and getting at least one 6?

$$1 - 0.343 = 0.657$$

(h) If you want a series of rolls to average between 3 and 5, and you get to choose the number of rolls before starting, should you choose 10 rolls or 100 rolls? Why?

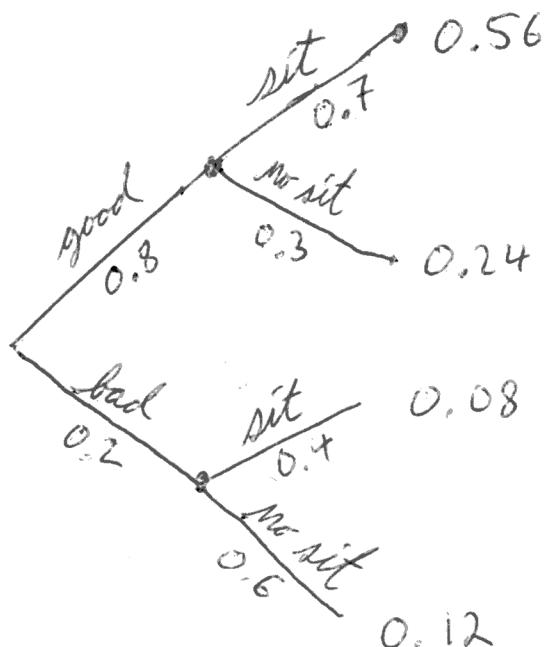
100 rolls will usually make sample mean $\approx \mu = 4$.

(i) If you want a series of rolls to average between 1 and 3, and you get to choose the number of rolls before starting, should you choose 10 rolls or 100 rolls? Why?

10 rolls, we want the variability of ~~small~~ small samples to help us get a sample mean far from the expected value (4).

Q7. (10 points) In a kennel, 80% of dogs are good dogs; the others are bad dogs. If a dog is good, it has a 70% chance of sitting on command. If a dog is bad, it has a 40% chance of sitting on command. You meet a dog, and it **does not** sit on command. What is the chance it is a good dog given it didn't sit on command?

(a) Draw a tree diagram.



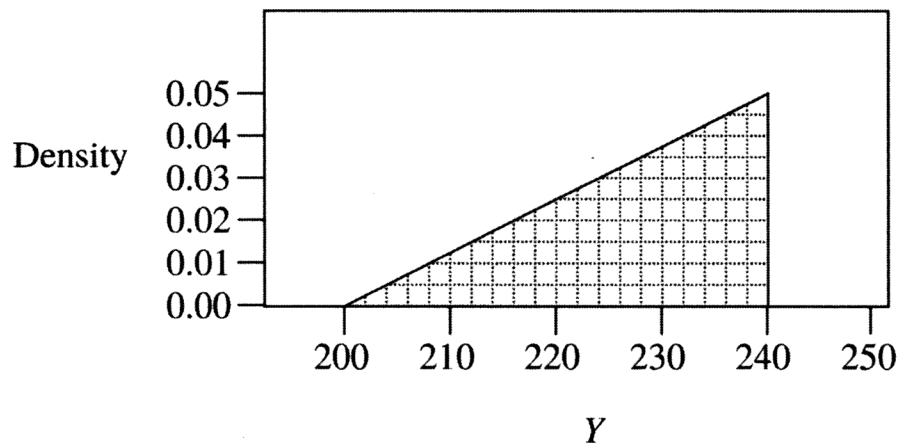
(b) Make a contingency table.

	good	bad	
sit	0.56	0.08	0.64
no sit	0.24	0.12	0.36
	0.8	0.2	

(c) Determine the probability the dog is good given it **did not** sit on command.

$$P(\text{good} \mid \text{no sit}) = \frac{0.24}{0.36} = \frac{2}{3} \approx \boxed{0.666}$$

- Q8.** (10 points) Let random variable Y be continuously distributed by the probability density function shown below. Notice the entire area is split into 100 percentile squares.



- (a) Evaluate $P(Y = 212)$. In other words, what is the probability that Y is exactly 212?

0

- (b) Evaluate $P(Y < 212)$. In other words, what is the probability that Y is less than 212?

0.09

- (c) Evaluate $P(220 < Y < 224)$. In other words, what is the probability that Y is between 220 and 224?

0.11

- (d) Estimate Q_1 .

220

- (e) Estimate the median.

228.1

- (f) Is the mean lower than, equal to, or greater than the median?

The mean is lower than the median.
(Left skew)