

Stat 241 Review Problems

1 Even when things are running smoothly, 5% of the parts produced by a certain manufacturing process are defective.

a) If you select 10 parts at random, what is the probability that none of them are defective?

Suppose you have a quality control procedure for testing parts to see if they are defective, but that the test procedure sometimes makes mistakes:

- If a part is good, it will fail the quality control test 10% of the time.
- 20% of the defective parts go undetected by the test.

b) What percentage of the parts will fail the quality control test?

c) If a part passes the quality control test, what is the probability that the part is defective?

d) The parts that fail inspection are sold as “seconds”. If you purchase a “second”, what is the probability that it is defective?

2 Numbers in a hat. In a hat are slips of paper with the number 1 – 100 on them.

a) If you reach in and grab one slip of paper, what is the probability that the number will be greater than 80?

b) If you reach in and grab one slip of paper, what is the probability that it is either greater than 80 or an odd number?

3 A manufacturing process produces parts that might have one of two faults. Let’s call them type 1 and type 2. 1% of parts have a type 1 fault, and 2% of faults have a type 2 fault.

a) If having the two types of faults are **independent** events, what is the probability that a part has a fault (of either type)?

b) If having the two types of faults are **mutually exclusive events**, what is the probability that a part has a fault (of either type)?

4 Lucky Number. The Lucky Number games works as follows. You pick a lucky number (1 – 6) and then roll three dice. If none of the dice match your lucky number, you get nothing. If one of the dice match your lucky number, you get \$3. If two of the dice match your lucky number, you get \$10. If all three of the dice match your lucky number, you get \$50.

a) Let X be your winnings in one play of the game. What is $E(X)$?

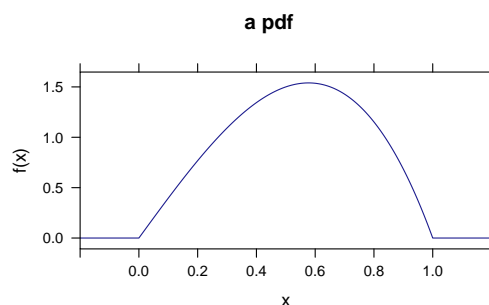
b) What does $E(X)$ tell you about this game?

5 The **pdf** for a continuous random variable X is

$$f(x) = \begin{cases} 4(x - x^3) & \text{when } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Determine $P(X \leq \frac{1}{2})$.

b) Determine the mean and variance of X .



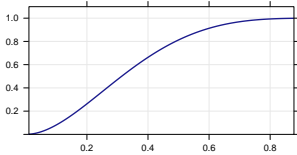
6 The **kernel** of a continuous distribution is given by $k(x) = 4 - x^2$ on the interval $[-2, 2]$.

- a) Determine the pdf of the distribution.
- b) Compute the mean and variance of the distribution.

7 The **cdf** for a random variable X is given by $F(x) = x^2$ on $[0, 1]$.

- a) What is $P(X \leq \frac{1}{2})$?
- b) What is the median of X ?
- c) What are the expected value and variance of X ?

8 Below is the graph of the **cdf** of a random variable Y .



- a) Use the graph to estimate $P(Y \leq 1/2)$
- b) Use the graph to estimate the median of Y .
- c) Sketch the pdf for Y .
- d) Which is larger, the median of Y or $E(Y)$?

9 The `mtcars` data frame has some information from Motor Trend about cars.

```
head(mtcars)
##           mpg cyl  disp  hp  drat   wt  qsec vs  am  gear carb
## Mazda RX4      21.0   6  160  110  3.90  2.620 16.46  0   1    4    4
## Mazda RX4 Wag  21.0   6  160  110  3.90  2.875 17.02  0   1    4    4
## Datsun 710     22.8   4  108   93  3.85  2.320 18.61  1   1    4    1
## Hornet 4 Drive  21.4   6  258  110  3.08  3.215 19.44  1   0    3    1
## Hornet Sportabout 18.7   8  360  175  3.15  3.440 17.02  0   0    3    2
## Valiant        18.1   6  225  105  2.76  3.460 20.22  1   0    3    1
```

Write down R commands to do the following:

- a) Compute the mean weight (`wt`) of these cars.
- b) Compute the mean weight of these cars, separately for 4-cylinder, 6-cylinder, and 8-cylinder vehicles. (How many different ways can you do this?)
- c) Create a scatter plot of fuel efficiency (`mpg`) vs. weight.
- d) Create side by side scatter plots of fuel efficiency (`mpg`) vs. weight separately for 4-cylinder, 6-cylinder, and 8-cylinder vehicles.

Stat 241 Review Problems

- e) Make up additional numerical and graphical summaries you could make from these data and write down the R commands you would use to create them.

10 Claire and Donald conduct an experiment to measure the strength of paper plates. Using the same data, Claire creates a 98% confidence interval for the mean paper plate strength and Donald creates a 95% confidence interval for the mean paper plate strength. Whose interval will be wider? Explain your reasoning.

11 Here are some terms we have used in class: statistic, parameter, sample, population, estimate, estimand, estimator, confidence interval, standard uncertainty, relative uncertainty, standard error, margin of error, critical value.

- Explain each of the terms in your own words. Where this makes sense, give (or make up) a specific example.
- Find pairs or groups of terms that are related somehow and explain how they are related. (Have fun with this, see how many relationships you can find.)

12 Beth and Sarah are bowlers. Bowling scores are discrete (whole numbers), but we can model them with a continuous distribution. As a first approximation, let's model their bowling scores as with normal distributions. Use the summary information below (based on their last 25 games) to answer some questions about their bowling performance, assuming that their performances are independent (a pretty good model for bowling, since players do not directly affect each other).

```
##   response bowler min  Q1 median  Q3 max mean  sd  n missing
## 1   score   Beth 109 154   175 196 241  175 31.8 25     0
## 2   score  Sarah 137 166   180 194 223  180 20.9 25     0
```

- Who is the more consistent bowler? How do you know?
- What is the probability that Sarah scores 200 or more?
- What is the probability that Beth scores 200 or more?
- If they bowl together as a team, what is the probability that their total score is 400 or more?
- If they bowl against each other, what is the probability that Beth will win (i.e., have a higher score than Sarah)?

13 A sample of 25 pieces of laminate used in the manufacture of circuit boards was tested for warpage (measured in mm) under specified conditions. The mean warpage for the sample was 2.42 mm and the sample standard deviation was 0.26 mm.

- Compute the standard error of the mean for this sample.
- Compute a 95% confidence interval for the mean warpage of all circuit boards under these conditions.

14 A 1996 study of Scotch pine specimens reported a 95% confidence interval for the mean elasticity of 14,500±1100 MPa for the mean modulus of elasticity obtained 1 minute after applying a certain load.

True or False. This means that the researchers believe that 95% of Scotch pine specimens will have an elasticity between 13,400 and 15,600 MPa.

Explain your reasoning.

15 Below is the summary output for a simple linear model for predicting the length of a child's foot from the width (both in cm).

```
msummary(model)

##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.8172     2.9381   3.341 0.00192 **
## width       1.6576     0.3262   5.081 1.1e-05 ***
##
## Residual standard error: 1.025 on 37 degrees of freedom
## Multiple R-squared:  0.411, Adjusted R-squared:  0.3951
## F-statistic: 25.82 on 1 and 37 DF,  p-value: 1.097e-05
```

- a) Give a 95% confidence interval for the slope in this model.
- b) What is the predicted response if the predictor variable has a value of 8.4 cm?
- c) Here is the data for David.

	name	birthmonth	birthyear	length	width	sex	biggerfoot	domhand
1	David	5	88	24.4	8.4	B	L	R

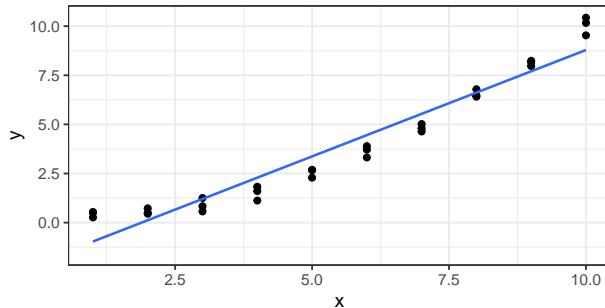
What is David's residual?

16 In linear modeling, what is the difference between a confidence interval and a prediction interval for the response? How do you get R to calculate them for you?

17 An experimenter measures the distance an object moves and how long it takes to move that far and records the measurements in the table below. Use this information to estimate the average speed of the object. Report your answer with the appropriate uncertainty to complete the table.

	quantity	estimate	uncertainty
	distance (m)	28.15	0.03
	time (sec)	6.51	0.02
	average speed (m/sec)		

18 The plot below shows some data that indicate a nonlinear relationship between y and x .



Sketch what the residual plot for this data will look like. Be sure to label the axes.