DATA 8

Week 8 – Hypothesis Testing, Probability

Slides by Suraj Rampure

Fall 2017

Administrative

Project 2 released Friday!

Once again, you should find partners within our lab – it'll make it much easier to work on it.

Midterm Grades Released!

You can submit regrade requests on Gradescope itself. If you have any issues with this, or are in any way concerned about your performance in the class, e-mail me and we can set up a time to chat.

Thanks to Vinitra, whose slides I used as reference when creating these this week!

Experimental vs. Theoretical

When we run an experiment and look at observations, we find an **empirical distribution**.

Flip a coin 100 times, see 53 heads and 47 tails.

The true values actually come from the **probability distribution**.

On each flip, a coin is equally likely to land heads or tails.

Parameters vs. Statistics

Characteristics of a population are called **population parameters**.

population mean, population variance, population standard deviation

Characteristics of a sample of a population are called **statistics**.

sample mean, sample variance, sample standard deviation

Bias and Variance

Bias is how much a statistic over-estimates or under-estimates a parameter.

Variance is a measure of how spread out values are – how much they "vary".

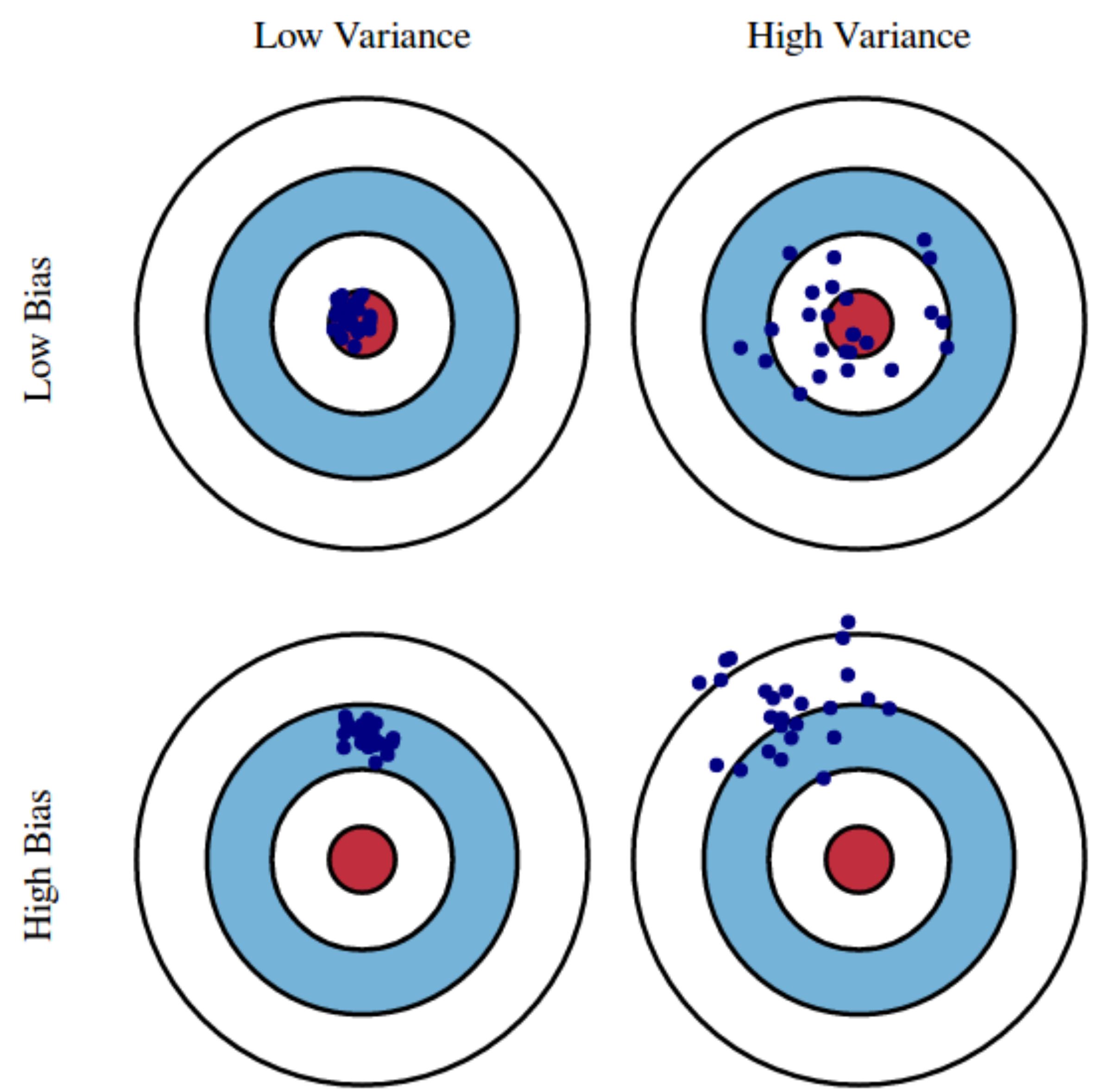


Fig. 1 Graphical illustration of bias and variance.

Source: https://i.stack.imgur.com/r7QFy.png

Hypothesis Testing

Null Hypothesis (H0): Observations are due to random chance

Alternative Hypothesis (H1): Something other than chance has influenced the observations

1. State null and alternative **hypotheses**
2. Define and compute a **test statistic** to help choose between hypotheses
3. The **probability distribution** of the test statistic
    1. What the test statistic might be if the null hypothesis were true
    2. Approximate the probability distribution by an empirical distribution
4. **Conclusion: Is the observed statistic consistent with the null distribution?**

Hypothesis Testing – Creating The Hypotheses

Consider the example of the rocket landing locations.

Alternative Hypothesis: This landing was special; its location was a draw from some other distribution, not the distribution from which the other 1100 landing locations were drawn.

Null Hypothesis: This landing was drawn from the same distribution as the other 1100 landing locations.

p-Values

The **p-value** is the **probability of the observation or something more extreme occurring under the null hypothesis**.

In our hypothesis test, using our test statistic, we're finding the probability of the observation occurring under the null. If we find that this probability is very low, we **reject the null hypothesis**.

We usually choose a **p-value cutoff** of 1% or 5%. The cutoff value is the probability of rejecting the null even though the data actually came from the null.

Putting it All Together

Let's say I flipped a coin (that I thought was fair) 100 times, and saw 65 heads and 35 tails. Suppose I wanted to test whether or not the coin was actually fair.

Null Hypothesis: The coin is fair, and any results are due to random chance.

Alternative Hypothesis: There is something other than random chance influencing outcomes – this coin is biased towards heads.

Putting it All Together

```
coin = [“H”, “T”]
outcomes = make_array()
for i in np.arange(10000):
 flips = np.random.choice(coin, 100)
 outcomes = np.append(outcomes, np.count_nonzero(flips == “H”))
p_value = np.count_nonzero(outcomes >= 65) / 10000
```

The value of `p_value` after running the above code will be the empirical probability of seeing 65 or more heads in 100 flips assuming the null hypothesis is true (assuming that the coin is fair).

Putting it All Together

If you run the previous code yourself, you'll see that `p_value` is usually some extremely small decimal value (~ 0.002). This means the chances of seeing 65 heads in 100 flips with a fair coin is under 0.2%, meaning that it is extremely unlikely that the coin is fair.

Since 0.2% is lower than any reasonable p-value cutoff (1% or 5%), we'd **reject the null hypothesis** in this case. This means we don't believe the coin is actually fair.

Probability – Extra Practice

With Replacement: After sampling, we are **able to choose these items again**

For example: Drawing 1 marble 5 times from a bag of 100 – if we do this with replacement, in each of the 5 samples we're drawing from 100 marbles)

Without Replacement: After sampling, we are **not able to choose these items again**

For example: Drawing 1 marble 5 times from a bag of 100 – if we do this without replacement, in the first sample we draw from 100 marbles, in the second we draw from 99, in the third from 98, and so on

1. Annie rolls a die 5 times. What is the chance that at least one of her rolls is a 5 or higher?

→ Either at least one of her rolls is ≥ 5, or none of them are

$\rightarrow P(\text{at least 1 roll} \geq 5) + P(\text{no rolls} \geq 5) = 1$

$P(\text{at least 1 roll} \geq 5) = 1 - \boxed{P(\text{no rolls} \geq 5)}$

on each roll, $P(\text{rolling} < 5) = \frac{4}{6}$

$\{\underbrace{1, 2, 3, 4}_{4}, 5, 6\}$

$$\rightarrow P(\text{at least 1 roll} \geq 5) = 1 - \underset{\text{roll 1}}{\left(\frac{4}{6}\right)}\underset{2}{\left(\frac{4}{6}\right)}\underset{3}{\left(\frac{4}{6}\right)}\underset{4}{\left(\frac{4}{6}\right)}\underset{5}{\left(\frac{4}{6}\right)}$$

$$= \boxed{1 - \left(\frac{4}{6}\right)^5}$$

2a. Nishant rolls a die 6 times. What are the chances that all 6 rolls are a 5?

a) $P(\text{all are } 5) = P(1^{st} \text{ is } 5) \cdot P(2^{nd} \text{ is } 5) \cdot \text{----} \cdot P(6^{th} \text{ is } 5)$

$= \left(\frac{1}{6}\right) \cdot \left(\frac{1}{6}\right) \cdot \text{----} \cdot \left(\frac{1}{6}\right)$

$= \left(\frac{1}{6}\right)^6$

2b. Nishant rolls a die 6 times. What are the chances that all 6 rolls are the same value?

b) $P(\text{all are the same}) = P(\text{all are } 1) + P(\text{all are } 2)$

$+ P(\text{all are } 3) + \dots + P(\text{all are } 6)$

$= 6 \cdot P(\text{all are } 5) = 6 \cdot \left(\frac{1}{6}\right)^6$

$= \boxed{\left(\frac{1}{6}\right)^5}$

ALTERNATIVELY → let the first roll be anything.

$P(\text{all are same}) = P(2^{nd} \text{ same as first}) \cdot P(3^{rd} \text{ same as first})$

$\cdot \dots \cdot P(6^{th} \text{ same as first})$

$= \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) \dots \left(\frac{1}{6}\right) = \boxed{\left(\frac{1}{6}\right)^5}$

2c. Nishant rolls a die 6 times. What are the chances that all 6 rolls are different values?

c)

$P(\text{all are different}) = P(1^{st} \text{ is unique}) \cdot P(2^{nd} \text{ is unique})$
$\cdot P(3^{rd} \text{ is unique}) \cdot P(4^{th} \text{ is unique})$
$\cdot P(5^{th} \text{ is unique}) \cdot P(6^{th} \text{ is unique})$

$$= \boxed{\left(\frac{6}{6}\right)\left(\frac{5}{6}\right)\left(\frac{4}{6}\right)\left(\frac{3}{6}\right)\left(\frac{2}{6}\right)\left(\frac{1}{6}\right)}$$

3. Fahad draws three cards from a deck at random and without replacement. Which of these events has the greater chance? Or are the chances equal?

A: He draws a heart, followed by a diamond, followed by a spade. ; B: Of the cards he draws, one is a heart, one a diamond, and one a spade.

3. Event B has a greater chance of occuring.

$$P(\text{event } A) = \left(\frac{13}{52}\right) \cdot \left(\frac{13}{51}\right) \cdot \left(\frac{13}{50}\right)$$

13 hearts, 52 cards remaining

13 diamonds, 51 cards remaining

13 spades, 50 cards remaining

$$P(\text{event } B) = \left(\frac{39}{52}\right) \cdot \left(\frac{26}{51}\right) \cdot \left(\frac{13}{50}\right)$$

can be any suit other than clubs

can be any suit other than clubs and card 1

must be a diff. suit than clubs, card 1 and card 2

4. Henry draws two cards from a deck, without replacement. a) What is the chance he draws two aces? b) What is the chance he draws no aces?

4.

a) $P(2 \text{ aces}) = \left(\frac{4}{52}\right) \cdot \left(\frac{3}{51}\right) = \underbrace{\frac{\binom{4}{2}}{\binom{52}{2}}}_{\text{ignore if unknown}}$

3 aces, 51 cards remaining

b) $P(\text{no aces}) = \left(\frac{48}{52}\right) \cdot \left(\frac{47}{51}\right)$

48 non-aces

47 cards remaining (no aces, and must be different than card 1)

4c. Henry draws two cards from a deck, without replacement. What is the chance that he draws exactly one ace?

c) P(exactly one ace)

Method 1

2 ways → first is ace, second isn't OR first isn't, second is

$$P(\text{exactly one ace}) = \underbrace{\left(\frac{4}{52}\right)}_{\text{4 aces}}\underbrace{\left(\frac{48}{52}\right)}_{\text{48 non-aces}} + \underbrace{\left(\frac{48}{52}\right)\left(\frac{4}{52}\right)}_{\text{same as first case}}$$

4c. Henry draws two cards from a deck, without replacement. What is the chance that he draws exactly one ace?

Method 2

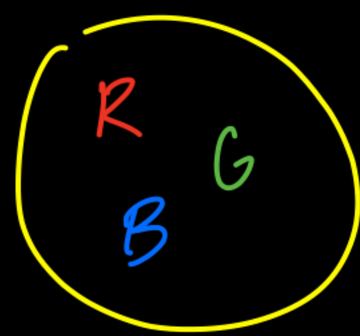
$$P(0 \text{ aces}) + P(\text{exactly } 1 \text{ ace}) + P(2 \text{ aces}) = 1$$

$$P(\text{exactly } 1 \text{ ace}) = 1 - P(0 \text{ aces}) - P(2 \text{ aces})$$

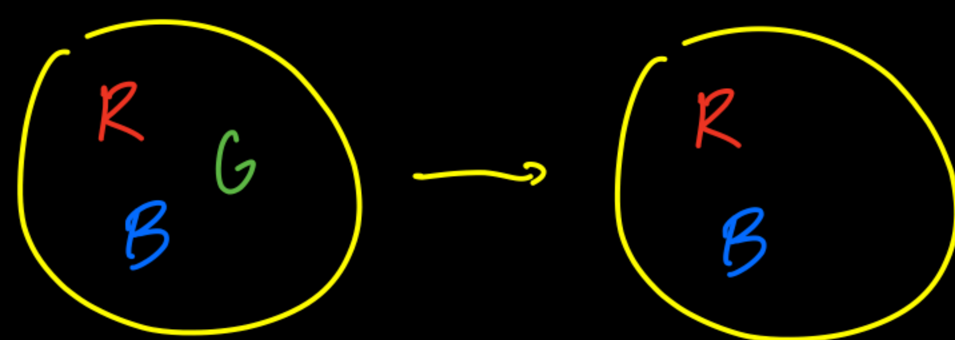
$$= 1 - \underbrace{\left(\frac{48}{52}\right)\left(\frac{47}{52}\right) - \left(\frac{4}{52}\right)\cdot\left(\frac{3}{51}\right)}_{\text{from parts a, b}}$$

5. Vinitra has a box with one red, one green and one blue ticket. She draws two tickets at random without replacement. a) What is the chance that the second ticket she draws is red, given that the first one is green? b) What is the chance that the second ticket she draws is red?

5.



a) $P(\text{second is green given first is red}) = \frac{1}{2}$



b) $P(\text{second ticket is red}) = P(\text{first isn't red})\, P(\text{second is red given first isn't red})$

$$= \left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = \boxed{\frac{1}{3}}$$

→ note: $P(1^{st} \text{ is red}) = P(2^{nd} \text{ is red}) = P(3^{rd} \text{ is red}) = \frac{1}{3}$. why?

6. Joseph gets ahold of the box, but he wants to draw tickets with replacement. Like Vinitra, he draws two tickets, but puts the first back in the box before drawing again. Fill in the rest of the table with the chance of each scenario, drawing with replacement (for Joseph), or without replacement (for Vinitra).

6.

	Joseph	Vinitra
P(First drawn is red)	1/3	1/3
P(Second drawn is red)	1/3	1/3
P(Second is red given first is green)	1/3	1/2

the fact that green was drawn first is irrelevant when we draw with replacement