Math 244 Lecture Notes

CHAPTER 18: CONFIDENCE INTERVALS FOR 1-PROPORTION

Overview: Today, we will look at how information about samples (**statistics**) can be used to make predictions about information about larger populations (**parameters**). To do this, we will briefly review a few terms and tools before diving head first into the topic of confidence intervals.

A **statistic** is a number that describes a sample. Examples of statistics include...

A **parameter** is a number that describes a population. Examples of parameters include...

We often want to know something about an entire population (e.g. what is the average salary for all Americans?) but it's not always possible to sample the entire population. In these cases, we rely on population models. Arguably the most important such model covered in MTH243 is the Normal Model/Distribution.

Normal Model The standard bell curve looks something like below. It is often a population model, so we describe the mean using ___ and the spread using ___.
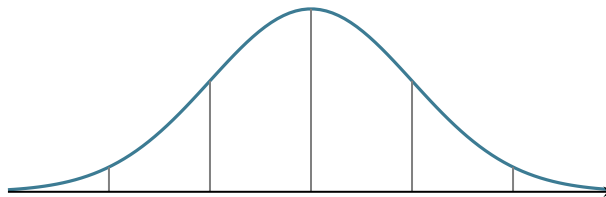
When using the normal distribution, we use the notation

$$X =$$

In our class, we will find the likelihood of events using the 68-95-99.7 Rule and with the aid of Z-scores.

Example 1. Suppose that $X = N(25, 5)$. Calculate $P(10 < X < 40)$.

Example 2. Suppose that $X = N(-3, 7)$. Calculate $P(-17 < x < 11)$.

What if I wanted to know $P(x < 10)$ in the previous example?

A Z-score tells us the number of standard deviations a value is away from the mean. We calculate it by doing...

This is a powerful tool for calculating probabilities using the normal distribution. For example, let's calculate $P(X < 10)$ using the following steps:

    (a) Draw a picture

    (b) Calculate the z-score(s)

    (c) Look up the z-score(s)

Example 3. The number of hours of sleep an average person in the US gets is normally distributed: $X = N(8, 0.5)$. Calculate $P(X < 9.25)$.

Example 4. Using the same information above, calculate $P(X > 7)$

Example 5. Using the same information above, calculate $P(7 < X < 9.25)$.

In MTH 243, we studied many different normal models used for making predictions:

- If we are studying sample means, we use the model...

$$\bar{X} = N($$

- If we are studying the number of success in n-trials, we use the model...

$$X = N($$

- If we are studying sample success rates/proportions, we use the model...

$$\hat{P} = N($$

We will revisit these models throughout the term, but lets start with the model for $\hat{P}$. As a reminder, the assumptions are...

(a) Random

(b) Independence

(c) Large n

Example 6. We are told that a coin is fair. With 68% certainty, what proportion of the time should we get heads in 100 flips?

Example 7. We are told that a coin is fair. With 95% certainty, what proportion of the time should we get heads in 100 flips?

> The interval $p \pm Z^* \sqrt{\frac{pq}{n}}$ gives us the interval that we expect $\hat{p}$ to be in based off population information.

This is the complete OPPOSITE of what we want. But...

> **Confidence Interval** The interval...
>
> gives us the interval that we expect $\hat{p}$ to be in based off the SAMPLE information. YEAH!!!

The assumptions are...

  (a) Random

  (b) Independence

  (c) Large n

Example 8. Using a random sample, an employee at the PCC bookstore finds that 97 out of 112 customers use their credit card to pay for books. Create a 95% confidence interval for what should go on for ALL PCC students.

Example 9. Let's create a 68% confidence interval for the proportion of students that identify as "female" using in-class data

A little terminology:

The **critical value** is denoted by z^*. It's the z-score corresponding to a specified confidence level.

The **standard error** is our guess for the SD based off the sample. $SE = \sqrt{\frac{\hat{p}\hat{q}}{n}}$

The **margin of error** is given by $MOE = z^* SE(\hat{p})$, or $ME = z^*\sqrt{\frac{\hat{p}\hat{q}}{n}}$.

The **confidence interval** for the true value of a proportion is determined by:

$$\text{best point estimate (center)} \ \pm \ \text{margin of error} \qquad \text{i.e.} \qquad \hat{p} \pm z^*\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Example 10. Using the current presidential approval information, let's determine the number of people sampled by Gallup. Assume 95% confidence: http://www.gallup.com/poll/113980/Gallup-Daily-Obama-Jo aspx

Example 11. What was the best point estimate in the previous problem?

Okay, but what about $C = 90\%$? What's Z^*?

Here's the general process...

- Place the area C in the center of a normal curve.
- Determine the area in BOTH tails, α. Then, determine the area in the left-tail.
- Use the table to get the z-score OR InvNorm(area). Switch the sign.

Example 12. Find Z^* for $C = 90\%$, $C = 95\%$, and $C = 99\%$.