Example 2: If the mean quiz score was 15.5 and standard deviation was 1.1, you were the 95^{th} percentile, what was your score? Reverse engineer the Z-score equation and use table.

$X = \mu + \sigma Z$

Question how do you find Z?

7 Lecture 8: Sampling Distribution, CLT, Binomial Approximation from Normal Distribution

Reality is wrong. Dreams are for real. – Tupac Shakur

7.1 Sampling Distribution

Example 1 It is known that the GPA **mean** at UCSC is $\mu = 3.22$ and **standard deviation** $\sigma = 0.15$. Imagine I ask 10 people their GPA and I record the mean and standard deviation from these 10 individuals. Imagine I do this a lot of times:

Groups	$\bar{x}$	s
Group 1	3.22	0.22
Group 2	3.89	0.15
Group 3	2.22	0.45
$\vdots$	$\vdots$	$\vdots$
Group 100,000	3.6	0.25

The means from each sample are random variables. Which makes sense, if we ask keep asking different samples we should get different means and standard deviations.

Which implies that the sample means can be model as a distribution. Which distribution do you think we will use?

What do you think the mean of this distribution will be?

BIG Concepts

1. The variability between the means will be smaller than the variability within each sample.
2. The sampling distributions of the mean allows us to determine the **probability** of a sample mean. One of the biggest reasons why we learned about probability is to discuss this idea.
   - Is the sample mean ________?
   - Is the sample mean not ________?

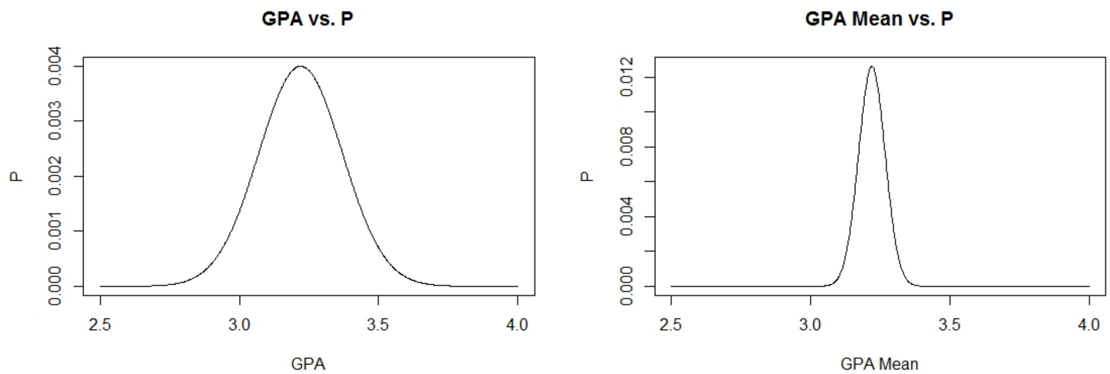


Figure 1: **Left Panel:** A random sample's distribution $(X_1, X_2, \cdots, X_n)$
Right Panel: The sample means' distribution $(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)$

When the sample size ________, the sample distribution of the sample mean becomes a normal distribution and becomes more and more ________.

The larger the sample the closer the sample mean will resemble the ________.

The sample mean will not vary as much as the sample size ________. Matter of fact, it well get closer to the true mean ________.

Here we are talking about the sample mean, but there are many other statistics

- **Sample Mean**
- Sample Median
- Sample Variance
- Sample Standard Deviation
- **Sample ___________**

7.2 Central Limit Theorem

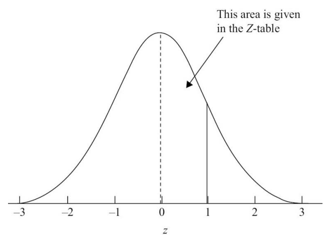
Based on these results we can conclude:

$$\bar{x} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}})$$

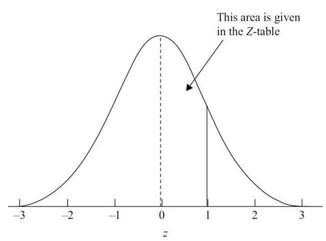
where __________ and __________

Example 1: Implementing the Sampling Distribution and CLT
Part 1: What is the probability that a random individual's GPA is less than 3.15? Part 2: What is the probability that a random sample of 10 mean GPA is less than 3.15? Part 3: What is the probability that a random sample of 100 mean GPA is less than 3.15?

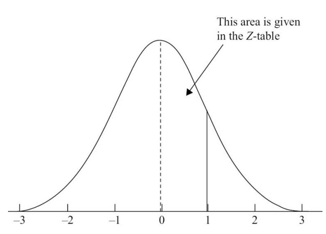
NOW LOOK
Random Individual $P(X \leq \_\_\_\_)$



Random Sample of size $n = 10$ $P(\bar{x} \leq \_\_\_\_)$



Random Sample of size $n = 100$ $P(\bar{x} \leq \_\_\_\_)$



7.3 Normal Approximation To the Binomial Distribution

It is the probability distribution of sample proportions, with all samples having the same sample size n.

Example 2: Use a random sample from the population, for example, $1,000$ people, to estimate the proportion who will vote for Bernie Sanders in the previous primary elections (using Polls).

IMPORTANT RESULT: Under certain conditions, the distribution of sample proportions approximates a normal distribution.

When working with a binomial distribution if ___ ≥ 5 and ___ ≥ 5 the binomial random variable has a probability distribution that can be approximated by a normal distribution, with mean and standard deviation:

$$\mu = np$$
$$\sigma = \sqrt{npq}$$

We must verify that it is reasonable to approximate the binomial distribution by the normal distribution

Using the same equation as the _______________

$$\hat{p} \sim N(\mu_{\hat{p}}, \sigma_{\hat{p}})$$

where $\mu_{\hat{p}} = np$ and $\sigma_{\hat{p}} = \sqrt{npq}$

Example 3: Let X denote the people out of the sample of 200 people who do not eat avocado.

Each person is independent within this sample, each person can either eat avocado or not, and the probability of eating avocado is $p = 0.5$ (will given to you directly or *indirectly on exam*) for everyone in the sample.

We can use the binomial distribution to model the number of individuals, X, who eat avocado from a sample size of 200.

Find the probability that less than 120 people like avocado.
Some Calculations preliminary = calculations that we can do

1. With $n = 200, p = 0.5$
   - $q = 1 - p = 0.5$
   - $np = 200(0.5) = 100 \quad (np \geq 5)$
   - $nq = 200(0.5) = 100 \quad (nq \geq 5)$
2. Then we can calculate:
   - $\mu = np = 200(0.5) = 100$
   - $\sigma = \sqrt{npq} = \sqrt{200(0.5)(0.5)} = 7.07$

The probability that less than 120 people like avocado $\implies P(X < 120)$