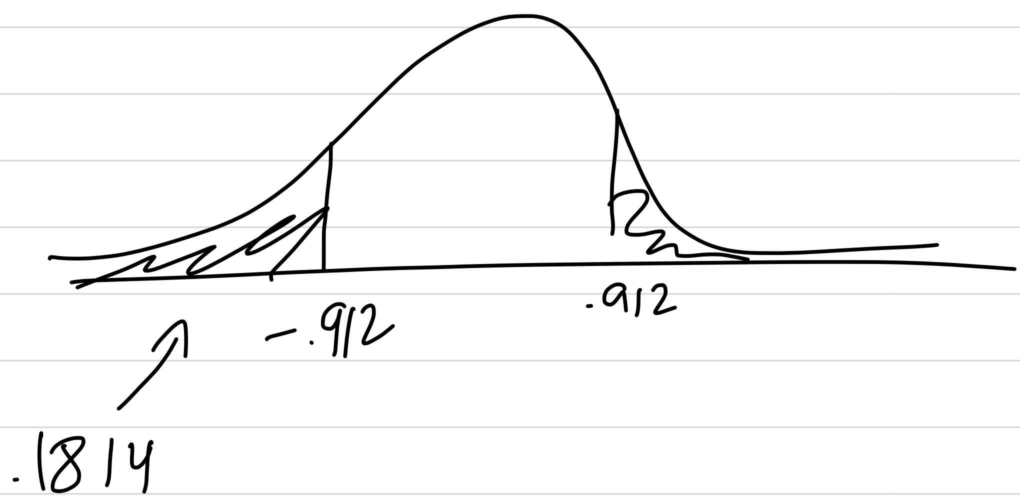
Ex: Let F be the random variable given by the number of fleas on a randomly selected household dog.

The distribution of F is not Normal, because it is discrete (b/c it only takes on integer values).

From studies, the population mean is approximately 2.7 with standard deviation 1.8. What is the approximate probability that a sample of 30 dogs will have a mean of more than 3?

By the Central Limit theorem, the distribution of $\bar{x}$ is approximately $N\left(2.7, \frac{1.8}{\sqrt{30}}\right) = N(2.7, .329)$

$$z = \frac{3-2.7}{.329} = .912$$



~18.14% chance of this sample mean being > 3

Chapter 16: Confidence Intervals

Statistical Inference for a Mean: we have an SRS, and the population is large compared to the sample size. We're measuring a variable whose distribution is $N(\mu, \sigma)$. We don't know μ, but we do know σ.

<u>Def</u>: A level C confidence interval for a parameter has two parts.

① An interval calculated from some data, of the form

estimate $\pm$ margin of error

② A confidence level C, which gives the probability that the interval will capture the true parameter value (i.e. the predicted success rate). The most common confidence level is 95%

What does this mean? For example, if you have a confidence interval of

$5 \pm .2$ with 95% confidence

We got to these numbers with a method that gives correct results 95% of the time.

2. $\frac{7.5}{\sqrt{654}}$

Population

μ unknown

$\sigma \approx 7.5$

$\xrightarrow{\text{SRS } n=654}$ $\bar{x} = 26.8 \pm .6$

$\xrightarrow{\text{SRS } n=654}$ $\bar{x} = 27.0 \pm .6$

$\xrightarrow{\text{SRS } n=654}$ $\bar{x} = 26.2 \pm .6$

$\vdots$

$$\left.\begin{array}{l} 26.8 \pm .6 \\ 27.0 \pm .6 \\ 26.2 \pm .6 \\ \vdots \end{array}\right\} \text{95\% will contain } \mu$$

95% of bands contain μ

Ex: A Gallup poll done in 2015 found that 26% of the 675 coffee drinkers in the sample were addicted to coffee. Here is how Gallup announced their results: "with 95% confidence, the maximum margin of error is ± 5 percentage points".

What is the confidence interval?

$26\% \pm 5\%$, so between 21% and 31%

What does this mean?

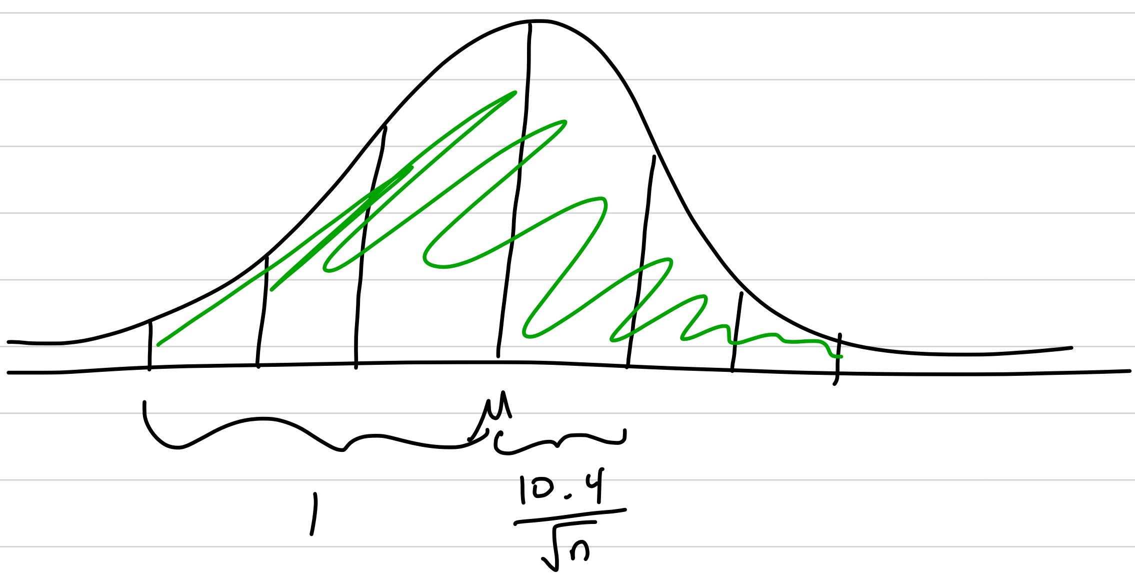
The chance that the actual proportion of the population addicted to coffee is between 21% and 31% is 95%

$N(\mu, 10.4)$

Sample of size n

distribution of $\bar{x}$ is $N\left(\mu, \frac{10.4}{\sqrt{n}}\right)$

$$\frac{10.4}{\sqrt{n}}$$



$$\frac{10.4}{\sqrt{n}} \cdot 3 = 1$$

distribution of $\bar{x}$ is $N(80, 4/\sqrt{25})$

$$= N(80, .8)$$

$$z = \frac{86-80}{.8} = 7.5$$

$$P(\text{positive} \mid \text{disease}) = \frac{P(\text{positive and disease})}{P(\text{disease})}$$

$$P(\text{disease}) = \frac{\#\text{ patients w/ disease}}{\#\text{ patients}}$$

$$= \frac{574}{1286}$$

$P(\text{positive and disease}) = \dfrac{564}{1286}$

$$\frac{564/1286}{574/1286} = \frac{564}{574} = 98.2\%$$

Priors

$P(A) = .26$

$P(B) = .49$

$P(M) = .2$

$P(D) = .05$

Posterior

$P(A \mid F)$

A = event of getting an associate degree

F = event of the recipient being female

$.61 = P(F \mid A)$

$P(F) = .5$

$$P(A \mid F) = \frac{P(F \mid A)\, P(A)}{P(F)}$$

$$= \frac{(.61)(.26)}{.5} = .317$$

$\sigma = 13$

$\bar{X}$: $n = 7$

$\bar{x}: \quad N\left(\mu, \frac{13}{\sqrt{7}}\right)$

Central Limit Theorem: sample of size n, the distribution of $\bar{x}$ is approximately $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.