# One-Sample Inference (Lectures 5 & 6)

BIOSTAT 201A Fall 2025

Discussion 3 – October 17, 2025

Cindy J. Pang

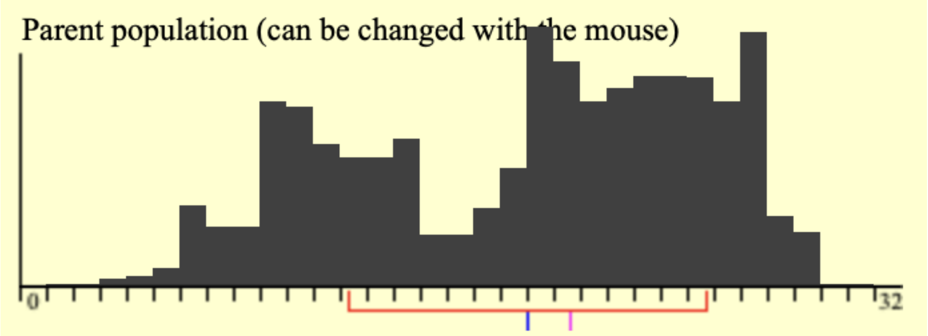
# Outline

1. Lecture 4. Point and Confidence Interval Estimation for Population Mean
    1. Sampling Distribution of the Mean
    2. Confidence Intervals for Means
2. Lecture 5. Point and Confidence Interval Estimation for Population Proportion
    1. Binomial Distribution
    2. Normal Approximation to the Binomial Distribution

# Outline

1. Lecture 4. Point and Confidence Interval Estimation for Population Mean
    1. Sampling Distribution of the Mean
    2. Confidence Intervals for Means
2. Lecture 5. Point and Confidence Interval Estimation for Population Proportion
    1. Binomial Distribution
    2. Normal Approximation to the Binomial Distribution

# 1.1 Sampling Distribution of the Mean




Population of units, X

$X \sim N(\mu, \sigma^2)$

Sample #1 of size n

Sample #2 of size n

...

Sample #k of size n

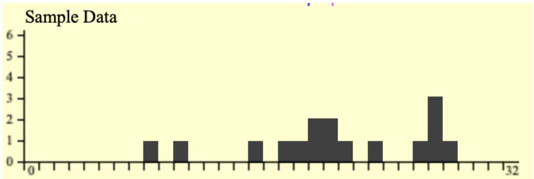
IMPT (for later)

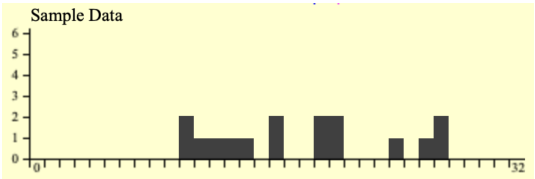
$$\boxed{SE(\bar{X}) = \sqrt{Var(\bar{X})} = \frac{\sigma}{\sqrt{n}}}$$

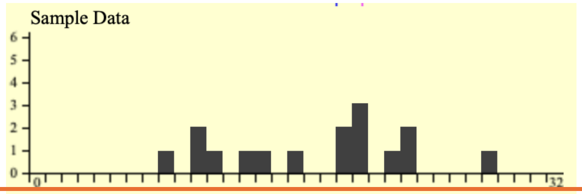
"Standard Error" of the Mean

$Var(\bar{X}) = \frac{\sigma^2}{n}$

$\mathbb{E}(\bar{X}) = \mu$




| | | | |
|---|---|---|---|
| Sample Means: | $\overline{x_1}$ | $\overline{x_2}$ | $\overline{x_k}$ |
| Sample Variances: | $s_1^2$ | $s_2^2$ | $s_k^2$ |

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

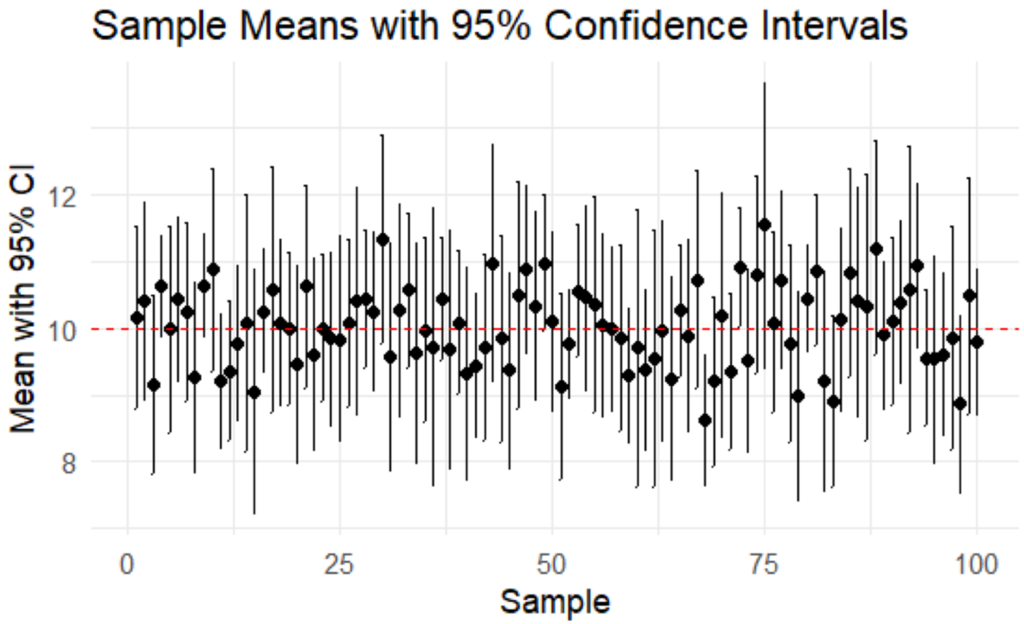
# 1.2 Confidence Intervals for Mean (Intuition)

$\mu = 10, \sigma^2 = 4$

| Samples from N(10,4) | | | | | |
|---|---|---|---|---|---|
| Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 | Sample 6 |
| 8.879049 | 12.448164 | 7.864353 | 10.852928 | 8.610586 | 10.506637 |
| 9.539645 | 10.719628 | 9.564050 | 9.409857 | 9.584165 | 9.942906 |
| 13.117417 | 10.801543 | 7.947991 | 11.790251 | 7.469207 | 9.914259 |
| 10.141017 | 10.221365 | 8.542217 | 11.756267 | 14.337912 | 12.737205 |
| 10.258575 | 8.888318 | 8.749921 | 11.643162 | 12.415924 | 9.548458 |
| 13.430130 | 13.573826 | 6.626613 | 11.377280 | 7.753783 | 13.032941 |
| 10.921832 | 10.995701 | 11.675574 | 11.107835 | 9.194230 | 6.902494 |
| 7.469877 | 6.066766 | 10.306746 | 9.876177 | 9.066689 | 11.169227 |
| 8.626294 | 11.402712 | 7.723726 | 9.388075 | 11.559930 | 10.247708 |
| 9.108676 | 9.054417 | 12.507630 | 9.239058 | 9.833262 | 10.431883 |

| Sample Means with 95% Confidence Intervals | | | | |
|---|---|---|---|---|
| Sample | Mean | Lower CI | Upper CI | Include Mean? |
| 1 | $\bar{x}_1 =$ 10.15 | 8.78 | 11.51 | 1 |
| 2 | $\bar{x}_2 =$ 10.42 | 8.93 | 11.90 | 1 |
| 3 | ⋮ 9.15 | 7.82 | 10.48 | 1 |
| 4 | 10.64 | 9.89 | 11.40 | 1 |
| 5 | 9.98 | 8.43 | 11.53 | 1 |
| 6 | $\bar{x}_6 =$ 10.44 | 9.22 | 11.67 | 1 |

95% of the confidence intervals will contain the true mean $\mu$




$\mu$

# 1.2 Confidence Intervals for Mean

Adjustment Factor

| Scenario | Assumptions | Confidence Interval for Mean |
|---|---|---|
| Known $\sigma^2$, sample size n > 30 | No Assumptions due to Central Limit Theorem! | $\boxed{\bar{x}} \pm \boxed{z_{1-\alpha/2}} \boxed{\frac{\sigma}{\sqrt{n}}}$ |
| Unknown $\sigma^2$ and large n (n > 200) | Population is (approximately) normal | $\boxed{\bar{x}} \pm \boxed{z_{1-\alpha/2}} \boxed{\frac{s}{\sqrt{n}}}$ |
| Unknown $\sigma^2$ and small n (n ≤ 200) | | $\boxed{\bar{x}} \pm \boxed{t_{n-1,1-\alpha/2}} \boxed{\frac{s}{\sqrt{n}}}$ |

Point Estimate of the Mean

Standard Errors of the Mean

**Exercise.**

## Pulmonary Disease

The data in Table 6.10 concern the mean triceps skin-fold thickness in a group of normal men and a group of men with chronic airflow limitation [5].

**TABLE 6.10 Triceps skin-fold thickness in normal men and men with chronic airflow limitation**

Unknown $\sigma^2$, small n

| Group | Mean | *sd* | *n* |
|---|---|---|---|
| Normal | $\bar{x}_N =$ 1.35 | $s_N =$ 0.5 | $n_N =$ 40 |
| Chronic airflow limitation | 0.92 | 0.4 | 32 |

*Source:* Adapted from *Chest*, *85*(6), 58S–59S, 1984.

| Scenario | Assumptions | Confidence Interval for Mean |
|---|---|---|
| Known $\sigma^2$, sample size n > 30 | No Assumptions due to Central Limit Theorem! | $\bar{x} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$ |
| Unknown $\sigma^2$ and large n (n > 200) | Population is (approximately) normal | $\bar{x} \pm z_{1-\alpha/2}\frac{s}{\sqrt{n}}$ |
| Unknown $\sigma^2$ and small n (n ≤ 200) | | $\bar{x} \pm t_{n-1,1-\alpha/2}\frac{s}{\sqrt{n}}$ |

a) Compute the Standard Error of the Mean for the Normal Group

$$SE(\bar{X}) = \frac{s}{\sqrt{n}} = \frac{0.5}{\sqrt{40}}$$

b) What is an Unbiased Estimator of the Population Mean?

$E(\bar{X}) = \mu \Rightarrow$ The sample mean

c) What is an Unbiased Estimator of the Population Variance?

$E(s^2) = \sigma^2 \Rightarrow$ The sample variance.

d) Compute a 95% Confidence Interval for the Mean of the Normal Group

$$\bar{x}_N \pm t_{n-1,\boxed{1-\alpha/2}}\frac{s}{\sqrt{n}} = 1.35 \pm t_{39,\,0.975}\left(\frac{0.5}{\sqrt{40}}\right)$$

$\alpha = 0.05$

$$= 1.35 \pm 2.02\left(\frac{0.5}{\sqrt{40}}\right)$$

$$= \boxed{[1.190, 1.510]}$$

e) Compute a 90% Confidence Interval for the Mean of the Normal Group

$$\bar{x}_N \pm \boxed{t_{n-1,1-\alpha/2}}\frac{s}{\sqrt{n}} = 1.35 \pm t_{39,0.95}\left(\frac{0.5}{\sqrt{40}}\right)$$

change this to $\alpha = .10$

$$= 1.35 \pm 1.68\left(\frac{0.5}{\sqrt{40}}\right)$$

$$= \boxed{[1.217, 1.483]}$$

f) Compare the 90% vs. 95% Confidence Intervals. Which one is larger, and what does that tell you about the precision of the estimate?

The 95% Confidence Interval is much wider compared to the 90% c.i, indicating that it is less precise.

# Outline

1. Lecture 4. Point and Confidence Interval Estimation for Population Mean
    1. Sampling Distribution of the Mean
    2. Confidence Intervals for Means
2. Lecture 5. Point and Confidence Interval Estimation for Population Proportion
    1. Binomial Distribution
    2. Normal Approximation to the Binomial Distribution

# 2.1 The Binomial Distribution

A **binomial random variable** counts the number of "successes" in the $n$ independent trials, where each trial has a probability $p$ of success

$$X \sim Binomial(n, p)$$

where $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ or "n-choose-k"

Conditions:

1. Experiment consists of *n* identical and independent (iid) trials
2. Dichotomous outcomes (only two possible outcomes) on each trial
3. Pr("success") = *p*, Pr("failure") = *1-p* where *p* is constant
4. Variable of interest is the number of successes observed during the *n* trials

# 2.1 The Binomial Distribution

A **binomial random variable** counts the number of “successes” in the $n$ independent trials, where each trial has a probability $p$ of success

$$X \sim Binomial(n, p)$$

where $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ or “n-choose-k”

**Exercise.** Let p = 0.2 and n = 5.

✓ use excel
=BINOM.DIST(n_success, n_trials, prob_success, FALSE)

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P(X=k)$ | 0.32768 | 0.4096 | 0.2048 | 0.0512 | 0.0064 | 0.00032 |

(a) What is the Expected Number of Successes?

$E(X) = 0(0.32768) + 1(.4096) + \ldots + 5(.00032) = \boxed{1}$ which is also $np = 5(0.2) = 1$

(b) What is the Variance of Successes?

$Var(X) = 0^2(0.32768) + 1^2(.4096) + \ldots + 5^2(.00032) - 1^2 = 1.8 - 1^2 = \boxed{0.8}$ is also $npq = \underbrace{5(.2)}_{=1}(.8) = 0.8$

(a) Compute the $\Pr(X \leq 3)$. $\Pr(X \leq 3) = \sum_{k=0}^{3} \Pr(X=k) = \Pr(X=0) + \ldots + \Pr(X=3) = \boxed{0.99328}$

# 2.2 Normal Approximation to the Binomial Distribution

| What you want to Estimate | Distribution | Assumption | Confidence Interval |
|---|---|---|---|
| Number of Successes | $X \sim N(np, np(1-p))$ | $np(1-p) \ \geq 5$ | |
| Proportions | $\hat{p} \sim N(p, \frac{p(1-p)}{n})$ where $\hat{p} = \frac{X}{n}$ | | $\hat{p} \pm z_{1-\alpha/2} \underbrace{\boxed{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}}_{SE(\hat{p})}$ |

**Exercise.** You obtain a simple random sample of 80 individuals, and 52 report having received a flu shot.

a) What is $\hat{p}$? $\hat{p} = \frac{X}{N} = \frac{52}{80} = \boxed{0.65}$

b) What is the standard error of $\hat{p}$? $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.65(1-0.65)}{80}} = \boxed{0.0533}$

c) What is the 95% confidence interval for $p$?

$z_{0.975} = 1.96 \Rightarrow 0.65 \pm 1.96\sqrt{\frac{0.65(1-0.65)}{80}} = \boxed{[0.545, 0.755]}$

d) Now suppose in a smaller nearby community, where you collect data from 80 individuals, and 12 have received a flu shot. Can you use the normal approximation to compute a 95% confidence interval? If not, explain why.

Yes, because $np(1-p) = 80\left(\frac{12}{80}\right)\left(1-\frac{12}{80}\right) = 12\left(1-{}^{12}/_{80}\right) = 10.2 \geq 5$.