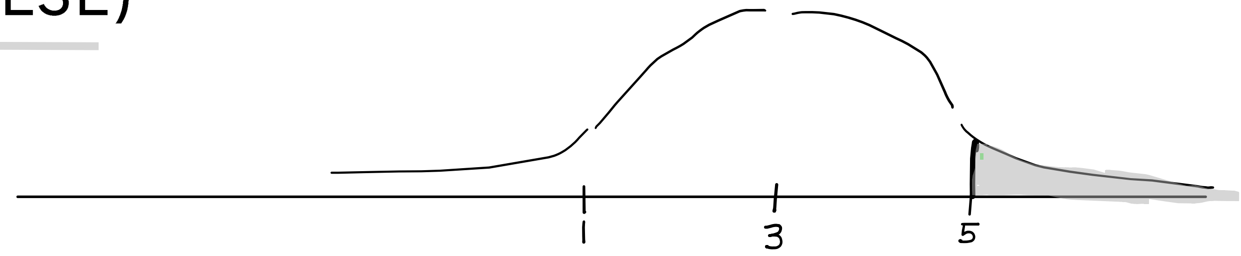
15 - Sampling Distribution (continued)

Parts of Normal Distributions

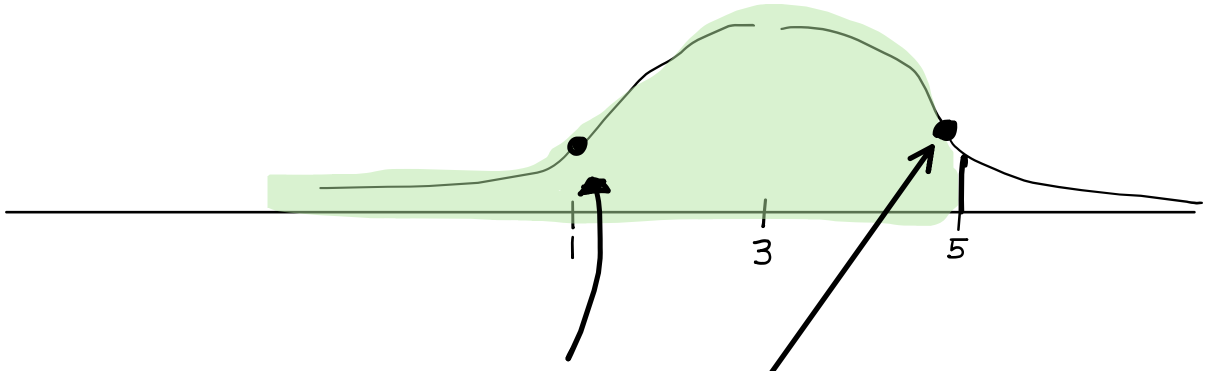
pnorm(5, mean=3, sd=2)

pnorm(5, mean=3, sd=2, lower.tail = FALSE)



Parts of Normal Distributions

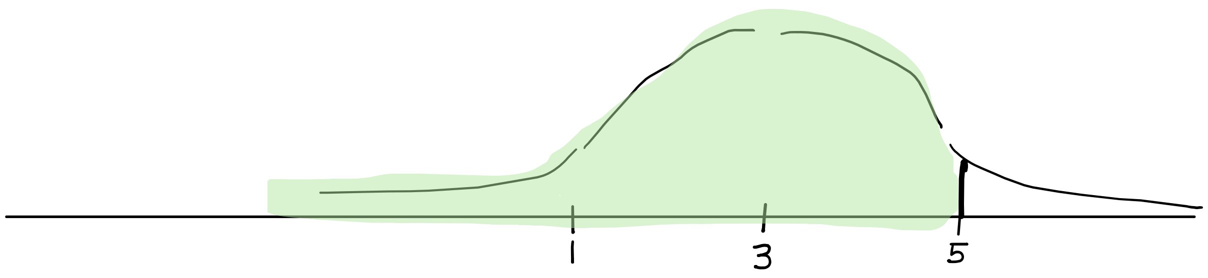
pnorm(5, mean=3, sd=2)



inflection point at 1 sd:
changes between curving up vs curving down

Parts of Normal Distributions

pnorm(5, mean=3, sd=2)

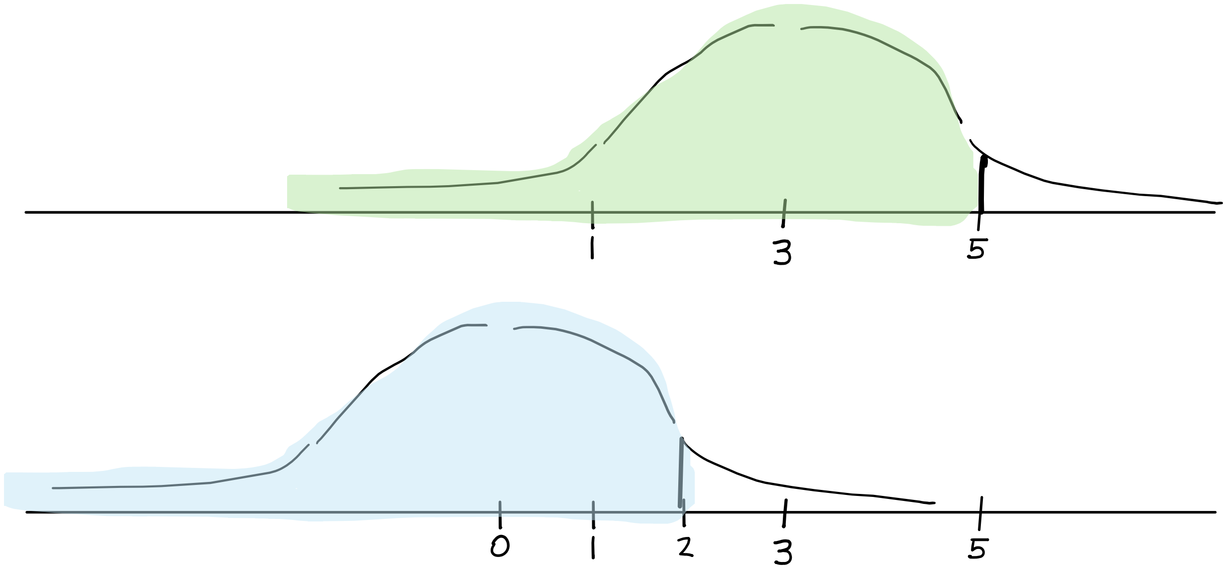


Parts of Normal Distributions

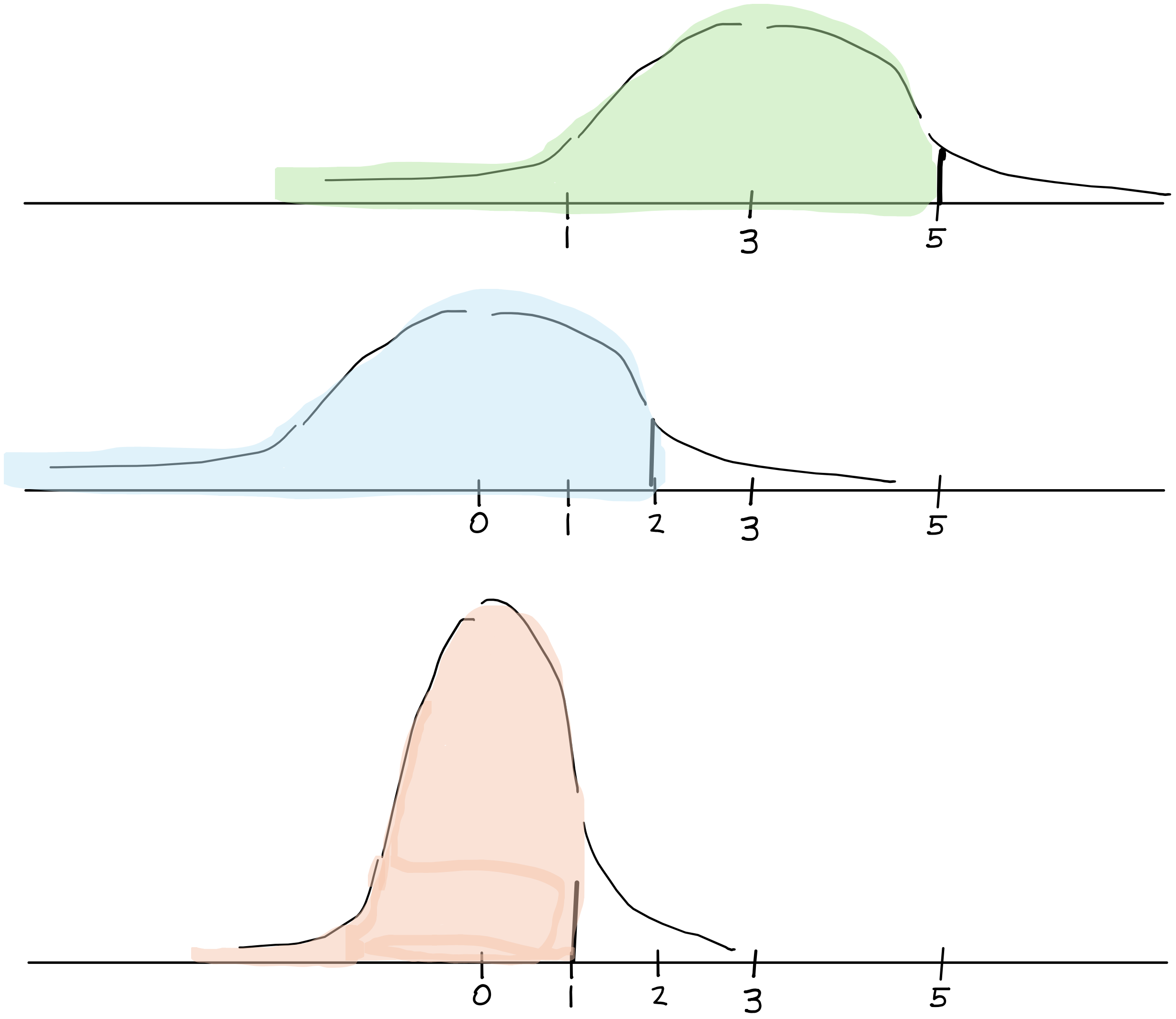
pnorm(5, mean=3, sd=2)

Subtract or translate by mean, $x \mapsto x - 3$:

pnorm(2, mean=0, sd=2)



Parts of Normal Distributions



pnorm(5, mean=3, sd=2)

Subtract or translate by mean, $x \mapsto x - 3$:

pnorm(2, mean=0, sd=2)

Divide or scale by standard deviation, $x \mapsto x/2$:

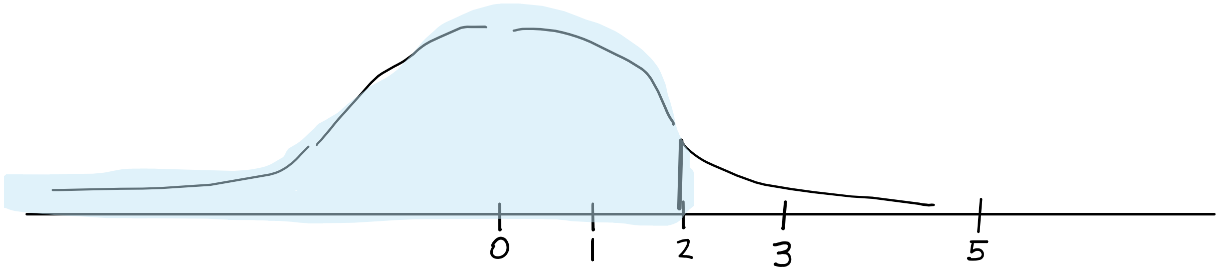
pnorm(1, mean=0, sd=1) = pnorm(1)

Parts of Normal Distributions

pnorm(5, mean=3, sd=2)

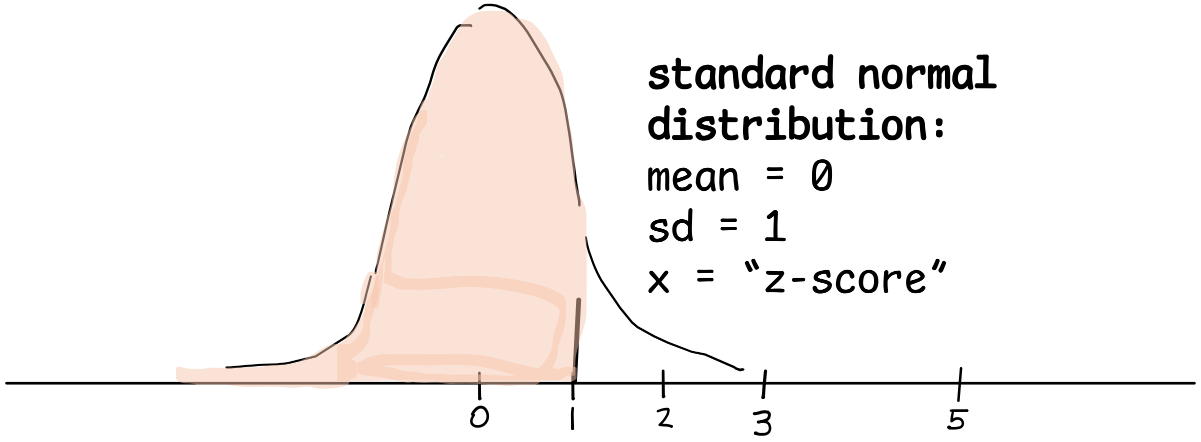
Subtract or translate by mean, $x \mapsto x - 3$:

pnorm(2, mean=0, sd=2)



Divide or scale by standard deviation, $x \mapsto x/2$:

pnorm(1, mean=0, sd=1) = pnorm(1)

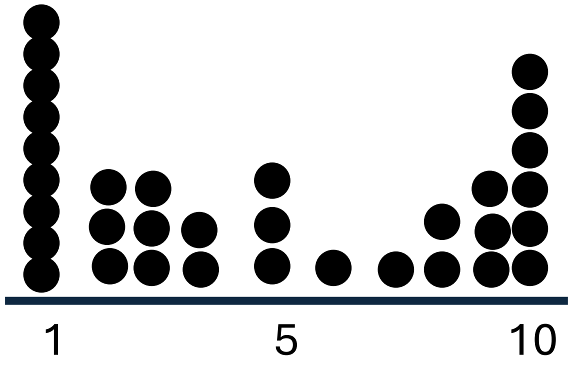


Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

Interview n=1 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



Interview n=2 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$

Interview n=10 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$

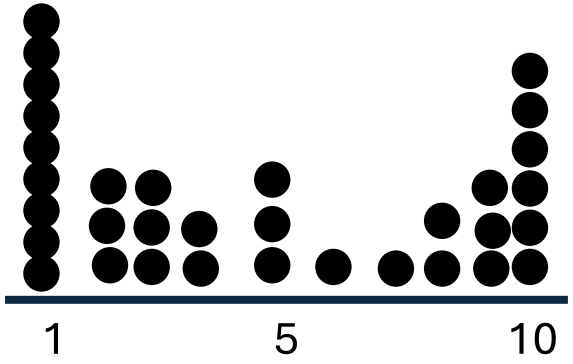
Interview n=200 person from N=200: **(census)**

Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

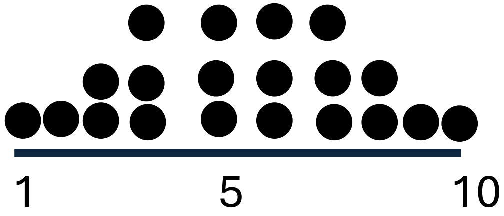
Interview n=1 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



Interview n=2 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$



Interview n=10 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$

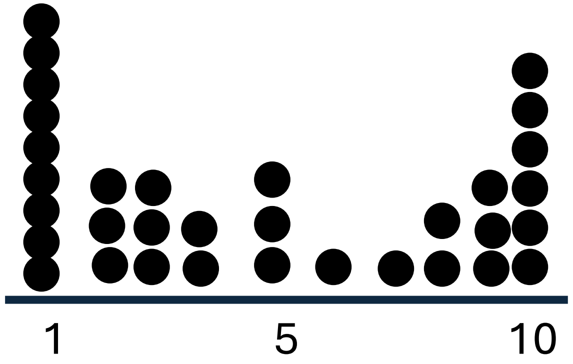
Interview n=200 person from N=200: **(census)**

Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

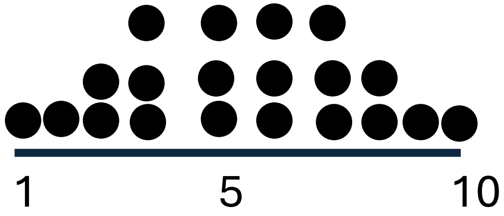
Interview n=1 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



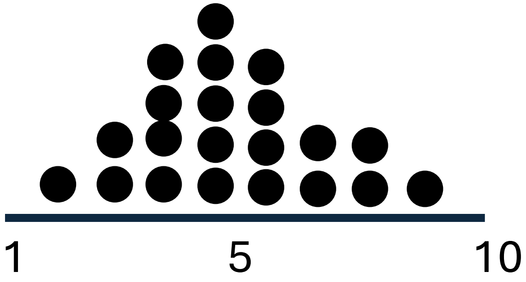
Interview n=2 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$



Interview n=10 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$



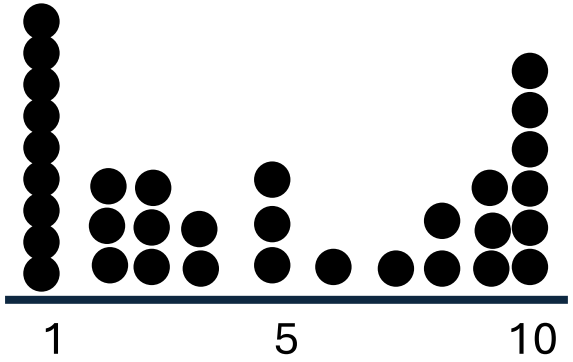
Interview n=200 person from N=200: **(census)**

Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

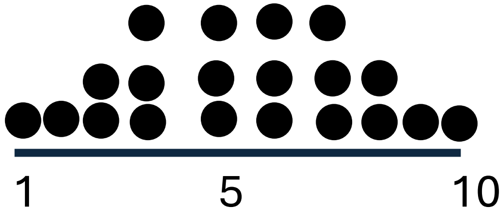
Interview n=1 person
from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



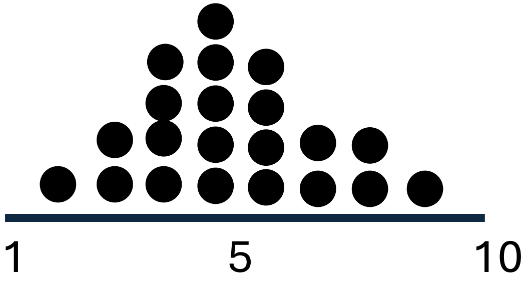
Interview n=2 person
from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$

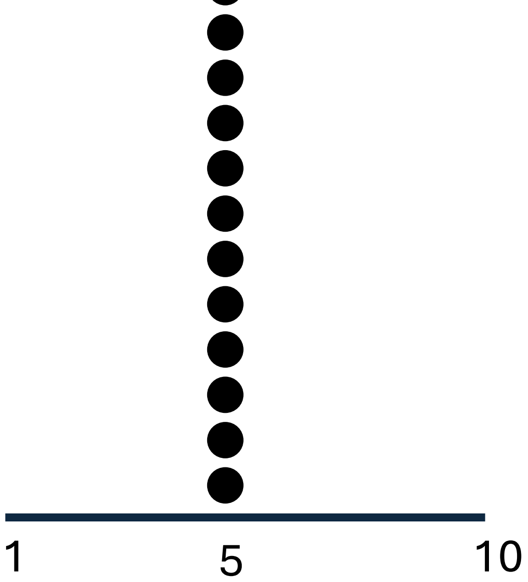


Interview n=10 person
from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$



Interview n=200 person
from N=200: **(census)**

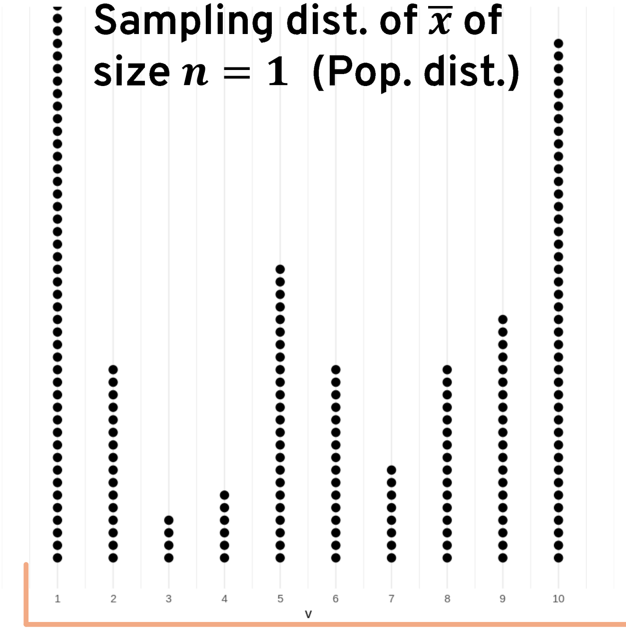


Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

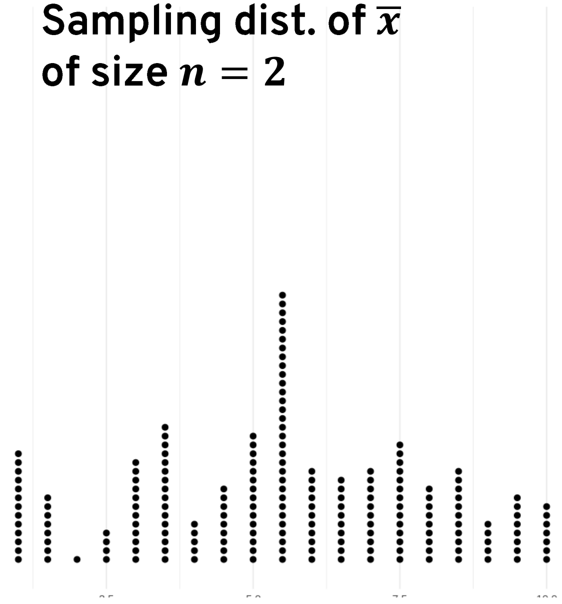
Interview n=1 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



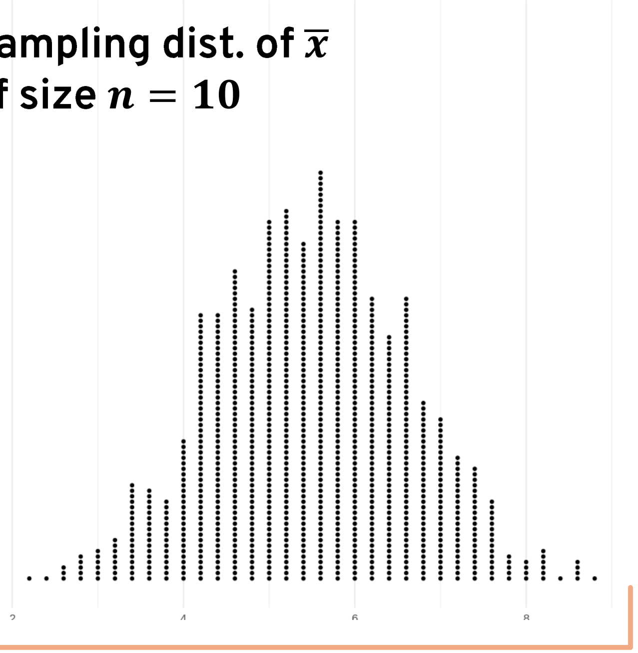
Interview n=2 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$

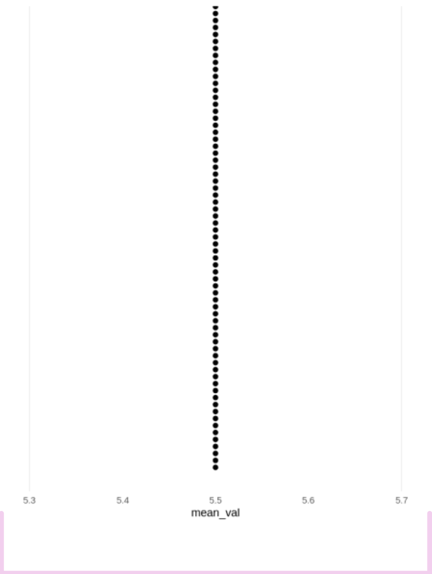


Interview n=10 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$



Interview n=200 person from N=200: **(census)**

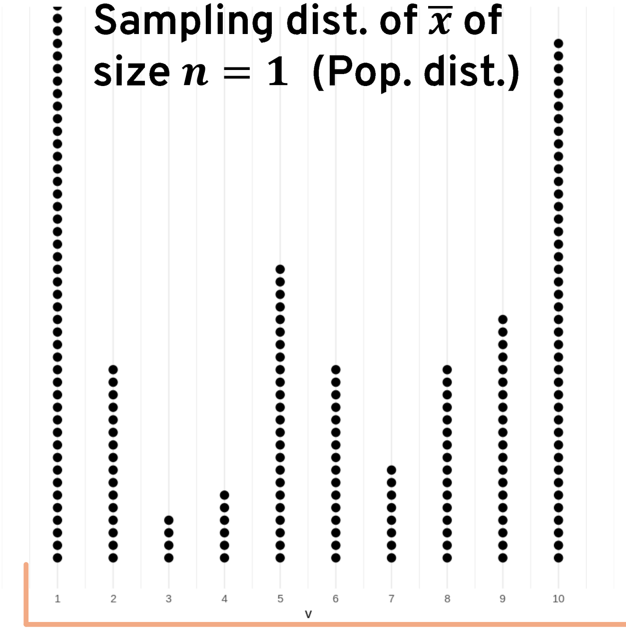


Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

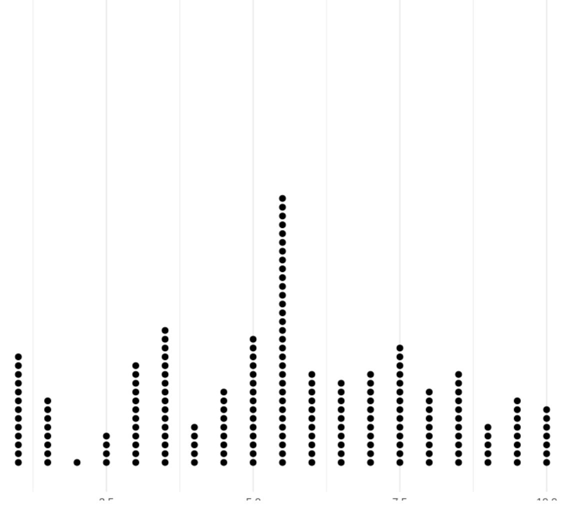
Interview n=1 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 1$ (Pop. dist.)



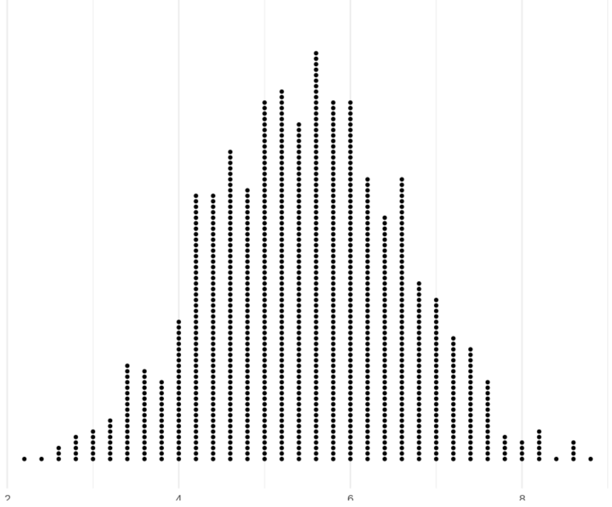
Interview n=2 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 2$

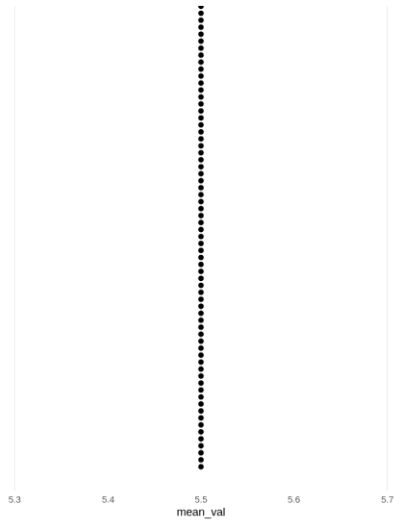


Interview n=10 person from N=200:

Sampling dist. of $\overline{x}$ of size $n = 10$



Interview n=200 person from N=200: **(census)**



Center: $\mu_{\bar{x}} = \mu$

Center: $\mu_{\bar{x}} = \mu$

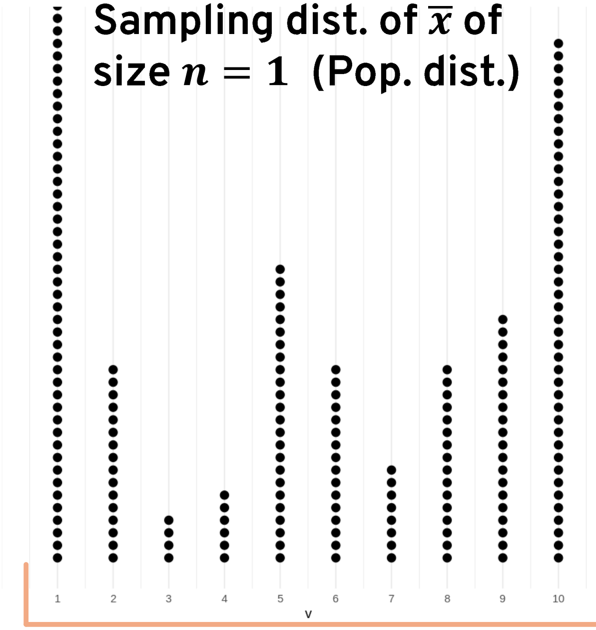
Spread: $\sigma_{\bar{x}} = 0$

Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

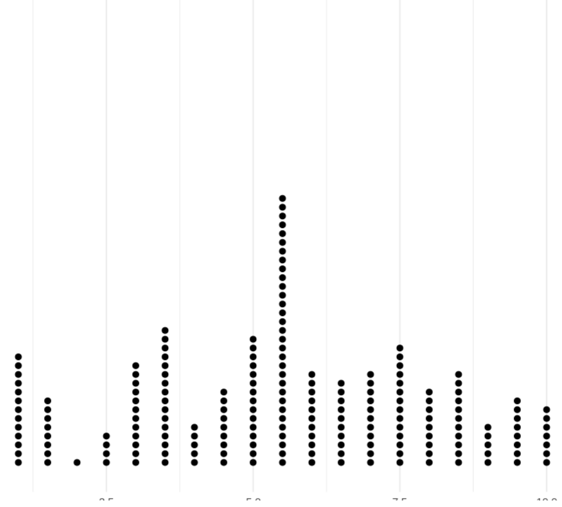
Interview n=1 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 1$ (Pop. dist.)



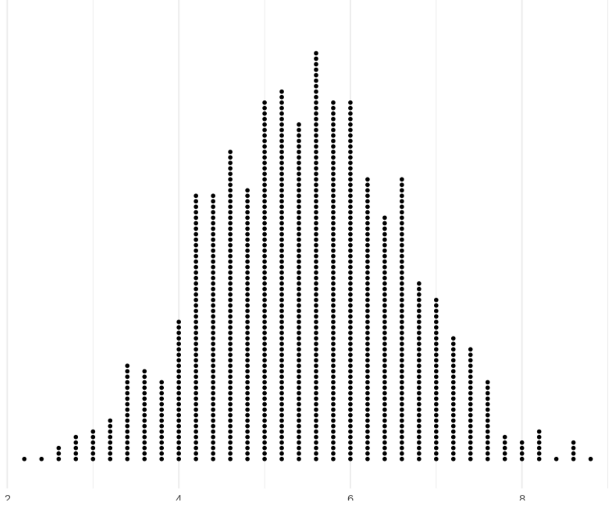
Interview n=2 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 2$

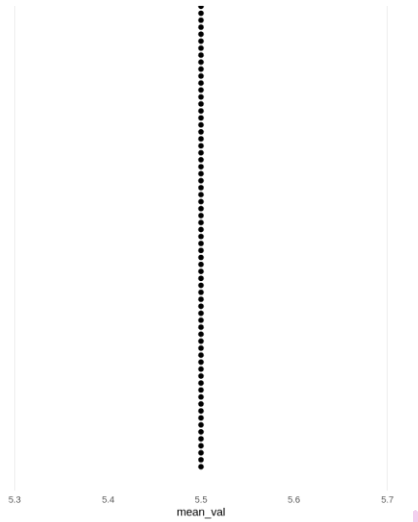


Interview n=10 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 10$



Interview n=200 person from N=200: **(census)**



Center: $\mu_{\bar{x}} = \mu$

Spread: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition)

Center: $\mu_{\bar{x}} = \mu$

Spread: $\sigma_{\bar{x}} = 0$

Sampling Distribution (Review)

Task: Rate 1-10 how much you like pineapple on pizza. Average respondent answers.

Interview n=1 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 1$ (Pop. dist.)

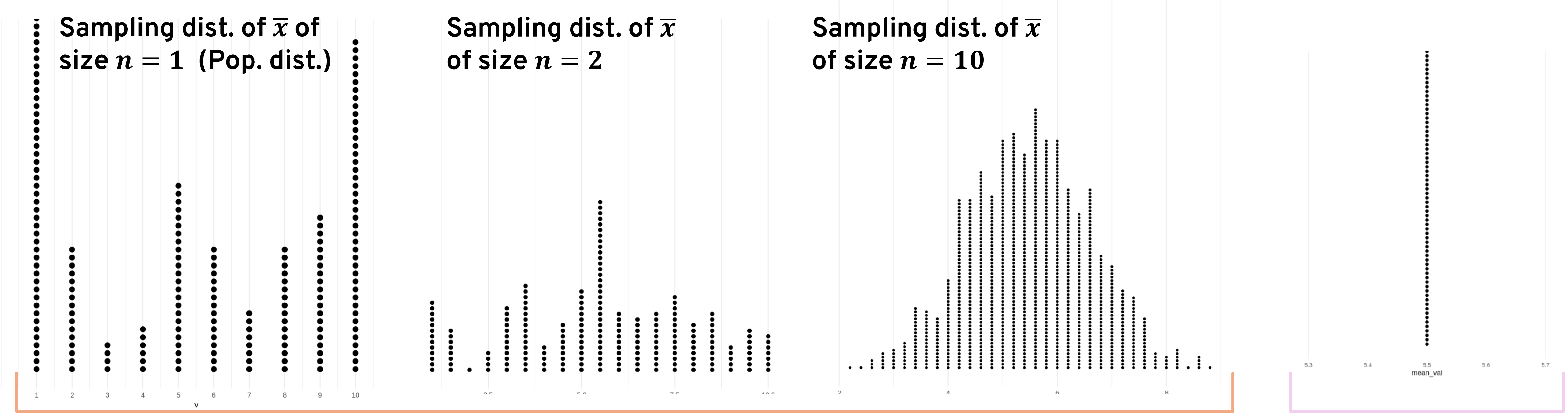
Interview n=2 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 2$

Interview n=10 person from N=200:

Sampling dist. of $\bar{x}$ of size $n = 10$

Interview n=200 person from N=200: **(census)**



Center: $\mu_{\bar{x}} = \mu$

Spread: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition)

Shape: Normal **if** pop. dist. is normal or **if** $n \geq 30$.

Center: $\mu_{\bar{x}} = \mu$

Spread: $\sigma_{\bar{x}} = 0$

Sampling Distribution of $\bar{x}$

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Sampling Distribution (Example 1)

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 1.
Sulfur odor threshold of adults follows a distribution with $\mu = 25$ and $\sigma = 7$. Find the mean and standard deviation the sampling distribution of $\bar{x}$ of size $n = 10$.

Sampling Distribution (Example 1)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 1.
Sulfur odor threshold of adults follows a distribution with $\mu = 25$ and $\sigma = 7$. Find the mean and standard deviation the sampling distribution of $\overline{\boldsymbol{x}}$ of size $n = 10$.

Answer.
Mean $\mu_{\bar{x}} = \mu = 25$

Sampling Distribution (Example 1)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).
- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 1.
Sulfur odor threshold of adults follows a distribution with $\mu = 25$ and $\sigma = 7$. Find the mean and standard deviation the sampling distribution of $\overline{x}$ of size $n = 10$.

Answer.
Mean $\mu_{\bar{x}} = \mu = 25$

Population has $\geq 10n = 10 \cdot 10 = 100$ adults so we can use sd formula:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7}{\sqrt{10}} \approx 2.2$$

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{\boldsymbol{x}}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{\boldsymbol{x}}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

```
pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE)  =
```

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{\boldsymbol{x}}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

```
pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE)  =
pnorm((66.5-64.5)/2.5, lower.tail=FALSE) = 0.211 = 21.1%
```

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE) =
pnorm((66.5-64.5)/2.5, lower.tail=FALSE) = 0.211 = 21.1%

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:
Mean: $\mu_{\bar{x}} = \mu = 64.5$.

Sampling Distribution (Example 2)

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

```
pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE)  =
pnorm((66.5-64.5)/2.5, lower.tail=FALSE) = 0.211 = 21.1%
```

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:
Mean: $\mu_{\bar{x}} = \mu = 64.5.$

Spread: Pop. has $N \geq 10n = 100$ young women so

$$\sigma_{\bar{x}} = \frac{2.5}{\sqrt{10}} \approx 0.79$$

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).
- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE) =
pnorm((66.5-64.5)/2.5, lower.tail=FALSE) = 0.211 = 21.1%

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:
Mean: $\mu_{\overline{x}} = \mu = 64.5.$

Spread: Pop. has $N \geq 10n = 100$ young women so

$$\sigma_{\overline{x}} = \frac{2.5}{\sqrt{10}} \approx 0.79$$

Shape: Pop. distribution is normal so sampling distribution is also approx. normal.

Sampling Distribution (Example 2)

Facts about sampling distribution of $\overline{\boldsymbol{x}}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 2.
Heights of young women are Normal with $\mu = 64.5$ and $\sigma = 2.5$.

(a) Probability a randomly selected young woman is taller than 66.5 inches:

pnorm(66.5, mean=64.5, sd = 2.5, lower.tail=FALSE) =
pnorm((66.5-64.5)/2.5, lower.tail=FALSE) = 0.211 = 21.1%

(b) Probability the mean height of an SRS of 10 young women exceeds 66.5 inches:
Mean: $\mu_{\bar{\boldsymbol{x}}} = \mu = 64.5.$

Spread: Pop. has $N \geq 10n = 100$ young women so

$$\sigma_{\bar{\boldsymbol{x}}} = \frac{2.5}{\sqrt{10}} \approx 0.79$$

Shape: Pop. distribution is normal so sampling distribution is also approx. normal.

Compute: $P(\bar{x} > 66.5) =$
$\text{pnorm}\left(\frac{66.5-64.5}{0.79}, \text{lower.tail=FALSE}\right) = 0.0057.$

Sampling Distribution (Example 3)

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Sampling Distribution (Example 3)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{\boldsymbol{x}}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 3.
Time for one technician to fix AC follows a right-skewed distribution with $\mu = 1$ hour, $\sigma = 1$ hour.
This week, your company services an SRS of 70 AC units in the city. You budget an average of 1.1 hours per unit for a technician to complete the work. Find probability the average maintenance time for 70 units exceeds 1.1 hours.

Sampling Distribution (Example 3)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 3.
Time for one technician to fix AC follows a right-skewed distribution with $\mu = 1$ hour, $\sigma = 1$ hour.
This week, your company services an SRS of 70 AC units in the city. You budget an average of 1.1 hours per unit for a technician to complete the work. Find probability the average maintenance time for 70 units exceeds 1.1 hours.

Answer. Sampling distribution of sample mean time has:
Mean: $\mu_{\bar{x}} = 1$ hour.

Sampling Distribution (Example 3)

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).
- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 3.
Time for one technician to fix AC follows a right-skewed distribution with $\mu = 1$ hour, $\sigma = 1$ hour.
This week, your company services an SRS of 70 AC units in the city. You budget an average of 1.1 hours per unit for a technician to complete the work. Find probability the average maintenance time for 70 units exceeds 1.1 hours.

Answer. Sampling distribution of sample mean time has:
Mean: $\mu_{\bar{x}} = 1$ hour.
Standard deviation: $\sigma_{\bar{x}} = \frac{1}{\sqrt{70}} \approx 0.12$ as there are $\geq 10 \cdot 70 = 700$ ACs in the city.

Sampling Distribution (Example 3)

Facts about sampling distribution of $\overline{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\overline{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).
- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 3.
Time for one technician to fix AC follows a right-skewed distribution with $\mu = 1$ hour, $\sigma = 1$ hour.
This week, your company services an SRS of 70 AC units in the city. You budget an average of 1.1 hours per unit for a technician to complete the work. Find probability the average maintenance time for 70 units exceeds 1.1 hours.

Answer. Sampling distribution of sample mean time has:
Mean: $\mu_{\overline{x}} = 1$ hour.
Standard deviation: $\sigma_{\overline{x}} = \frac{1}{\sqrt{70}} \approx 0.12$ as there are $\geq 10 \cdot 70 = 700$ ACs in the city.
Shape: Distribution is approx. normal as $n = 70 \geq 30$.

Sampling Distribution (Example 3)

Facts about sampling distribution of $\bar{x}$.
Population has size N, mean μ, standard deviation σ.
Take SRS of size n.
Then sampling distribution of sample mean $\bar{x}$ of size n has:

- Mean $\mu_{\bar{x}} = \mu$
- Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ if $N \geq 10n$ (10% Condition).

- Normal if population distribution is normal.
- Normal if population distribution is not too skewed and $n \geq 30$.

Example 3.
Time for one technician to fix AC follows a right-skewed distribution with $\mu = 1$ hour, $\sigma = 1$ hour.
This week, your company services an SRS of 70 AC units in the city. You budget an average of 1.1 hours per unit for a technician to complete the work. Find probability the average maintenance time for 70 units exceeds 1.1 hours.

Answer. Sampling distribution of sample mean time has:
Mean: $\mu_{\bar{x}} = 1$ hour.
Standard deviation: $\sigma_{\bar{x}} = \frac{1}{\sqrt{70}} \approx 0.12$ as there are $\geq 10 \cdot 70 = 700$ ACs in the city.
Shape: Distribution is approx. normal as $n = 70 \geq 30$.

Compute:

$$P(\bar{x} > 1.1) = \text{pnorm}\left(\frac{1.1 - 1}{0.2}, \text{lower.tail=FALSE}\right) = 0.308.$$

Sampling Distribution of $\hat{p}$

Idea: Proportion $\hat{\boldsymbol{p}}$ of people in a sample of size n that answer Yes to a question is a random variable:

Model: Each person has same probability p of answering Y (like a biased coin.) If 10% Condition holds:

Proportion of Y's in sample of size n is binomial distribution with $\mu_{\hat{\boldsymbol{p}}} = p$ and $\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{p(1-p)}{n}}$.

Sampling Distribution of $\hat{p}$

Idea: Proportion $\hat{\boldsymbol{p}}$ of people in a sample of size n that answer Yes to a question is a random variable:

Model: Each person has same probability p of answering Y (like a biased coin.) If 10% Condition holds:

Proportion of Y's in sample of size n is binomial distribution with $\mu_{\hat{\boldsymbol{p}}} = p$ and $\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{p(1-p)}{n}}$.

When is Sampling Distribution of $\hat{p}$ normal?

$p \approx 0$
n small

$p \approx 0.5$
n small

$p \approx 1$
n small

$p \approx 0$
n big

$p \approx 0.5$
n big

$p \approx 1$
n big

Sampling Distribution of $\hat{p}$

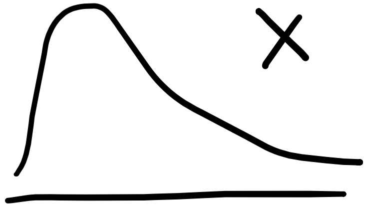
Idea: Proportion $\hat{\boldsymbol{p}}$ of people in a sample of size n that answer Yes to a question is a random variable:

Model: Each person has same probability p of answering Y (like a biased coin.) If 10% Condition holds:

Proportion of Y's in sample of size n is binomial distribution with $\mu_{\hat{\boldsymbol{p}}} = p$ and $\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{p(1-p)}{n}}$.

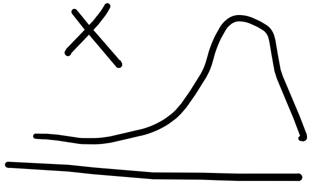
When is Sampling Distribution of $\hat{p}$ normal?

$p \approx 0$
n small



$p \approx 0.5$
n small

$p \approx 1$
n small



sd big

$p \approx 0$
n big

$p \approx 0.5$
n big

$p \approx 1$
n big

sd small

Sampling Distribution of $\hat{p}$

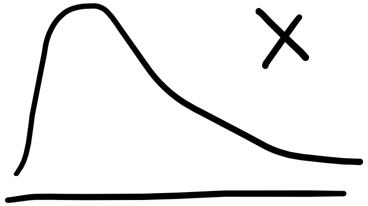
Idea: Proportion $\hat{\boldsymbol{p}}$ of people in a sample of size n that answer Yes to a question is a random variable:

Model: Each person has same probability p of answering Y (like a biased coin.) If 10% Condition holds:

Proportion of Y's in sample of size n is binomial distribution with $\mu_{\hat{\boldsymbol{p}}} = p$ and $\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{p(1-p)}{n}}$.

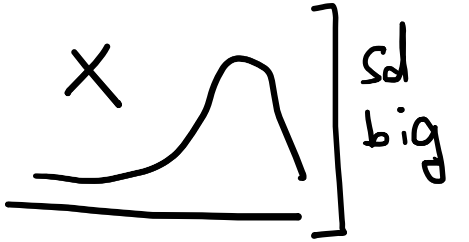
When is Sampling Distribution of $\hat{p}$ normal?

$p \approx 0$
n small

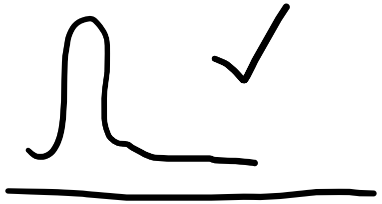


$p \approx 0.5$
n small

$p \approx 1$
n small

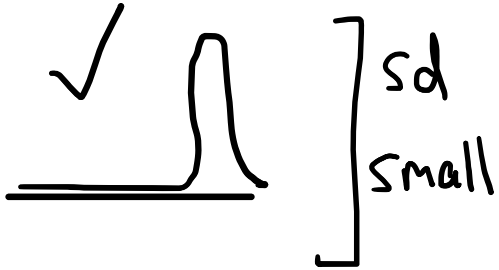


$p \approx 0$
n big



$p \approx 0.5$
n big

$p \approx 1$
n big



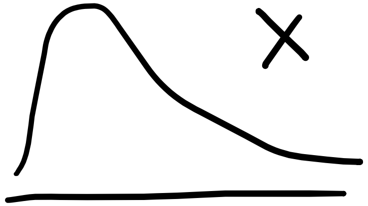
Sampling Distribution of $\hat{p}$

Idea: Proportion $\hat{\boldsymbol{p}}$ of people in a sample of size n that answer Yes to a question is a random variable:

Model: Each person has same probability p of answering Y (like a biased coin.) If 10% Condition holds:
Proportion of Y's in sample of size n is binomial distribution with $\mu_{\hat{\mathbf{p}}} = p$ and $\sigma_{\hat{\mathbf{p}}} = \sqrt{\frac{p(1-p)}{n}}$.

When is Sampling Distribution of $\hat{p}$ normal? **Answer: $\boldsymbol{np} \geq \mathbf{10}$ and $\boldsymbol{n(1-p)} \geq \mathbf{10}$ (Large Counts Condition)**

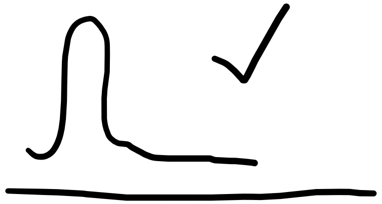
$p \approx 0$
n small



$p \approx 0.5$
n small

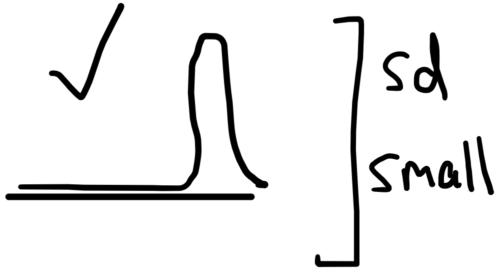
$p \approx 1$
n small

$p \approx 0$
n big



$p \approx 0.5$
n big

$p \approx 1$
n big



Sampling Distribution of $\hat{p}$

Facts about sampling distribution of $\hat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\hat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Sampling Distribution (Example 4)

Facts about sampling distribution of $\hat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\hat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Sampling Distribution (Example 4)

Facts about sampling distribution of $\hat{p}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{p}$ of size n has:

- Mean $\mu_{\hat{p}} = p$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Mean: $\mu_{\hat{p}} = 0.35.$

Sampling Distribution (Example 4)

Facts about sampling distribution of $\hat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\hat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Mean: $\mu_{\hat{\boldsymbol{p}}} = 0.35.$

Standard deviation: There are $\geq 10 \cdot 1500$ such students:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.35(0.65)}{1500}} \approx 0.0123$$

Sampling Distribution (Example 4)

Facts about sampling distribution of $\hat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\hat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Mean: $\mu_{\hat{\boldsymbol{p}}} = 0.35.$

Standard deviation: There are $\geq 10 \cdot 1500$ such students:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.35(0.65)}{1500}} \approx 0.0123$$

Shape: Approx. normal bc $np = 525, n(1-p) = 975 \geq 10.$

Sampling Distribution (Example 4)

Facts about sampling distribution of $\hat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\hat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\hat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Mean: $\mu_{\hat{\boldsymbol{p}}} = 0.35.$

Standard deviation: There are $\geq 10 \cdot 1500$ such students:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.35(0.65)}{1500}} \approx 0.0123$$

Shape: Approx. normal bc $np = 525, n(1-p) = 975 \geq 10.$

Compute: $P(0.33 \leq \hat{p} \leq 0.37)$

$$= \text{pnorm}\left(\frac{0.37 - 0.35}{0.0123}\right) - \text{pnorm}\left(\frac{0.33 - 0.35}{0.0123}\right) = 0.896.$$

Sampling Distribution (Example 4)

Facts about sampling distribution of $\widehat{\boldsymbol{p}}$.
Population has size N with proportion p of success. Take SRS of size n.

Then sampling distribution of $\widehat{\boldsymbol{p}}$ of size n has:

- Mean $\mu_{\widehat{\boldsymbol{p}}} = \boldsymbol{p}$
- Standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ if $N \geq 10n$ (10% Condition).
- Approx. Normal if $np \geq 10$ and $n(1-p) \geq 10$ (Large Counts Condition).

Example 4.
In an SRS of 1500 first-year college students, 35% of all first-year students attend college within 50 miles of home. Find the probability that the random sample of 1500 students will give a result within 2% of the true value.

Mean: $\mu_{\widehat{\boldsymbol{p}}} = 0.35$.

Standard deviation: There are $\geq 10 \cdot 1500$ such students:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.35(0.65)}{1500}} \approx 0.0123$$

Shape: Approx. normal bc $np = 525, n(1-p) = 975 \geq 10$.

Compute: $P(0.33 \leq \hat{p} \leq 0.37)$

$$= \text{pnorm}\left(\frac{0.37 - 0.35}{0.0123}\right) - \text{pnorm}\left(\frac{0.33 - 0.35}{0.0123}\right) = 0.896.$$

Conclude: About 90% of SRSs of size 1500 are within 2% of true proportion.