lecture4: random sampling

6 Mart 2025 Perşembe 10:33

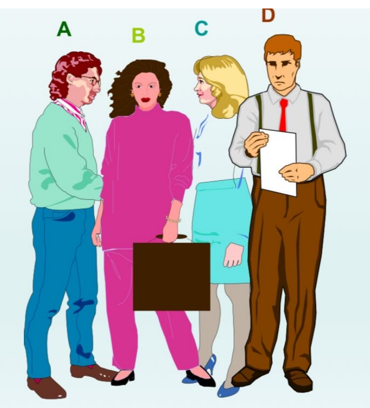
Population mean (μ) pop. variance (σ^2), pop. size (N), standard deviation (σ)

↳ Population characteristics

Sample mean $(\bar{x})$ sample variance (s^2), sample size (n) standard deviation (s)

↳ sample characteristics

- Population size N=4
- Random variable, X, is age of individuals
- Values of X: 18, 20, 22, 24 (years)



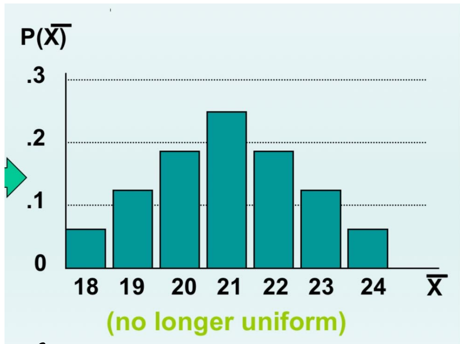
$$\mu = \frac{\sum x}{N} = \frac{18+20+22+24}{4} = 21$$

1st Obs	2nd Observation			
	18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

↳ samples of size n=2

1st Obs	2nd Observation			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

↳ sample means



↳ sample distribution

$$\boxed{\text{standard error of the mean} = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

sample size: n=16

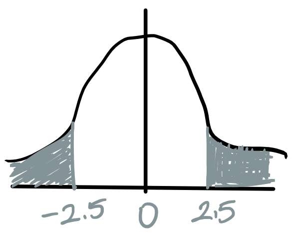
population X = lifespan of bulbs

$X \sim N(\mu=800, \sigma^2=(40)^2) \Rightarrow \bar{X} \sim N(\mu=800, \sigma_{\bar{x}}^2 = \frac{40^2}{16})$

$\sigma_{\bar{x}} = \frac{40}{\sqrt{16}} = 10$

$P(\bar{X} < 775) = ?$

$P\left(\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{775-800}{10}\right)$

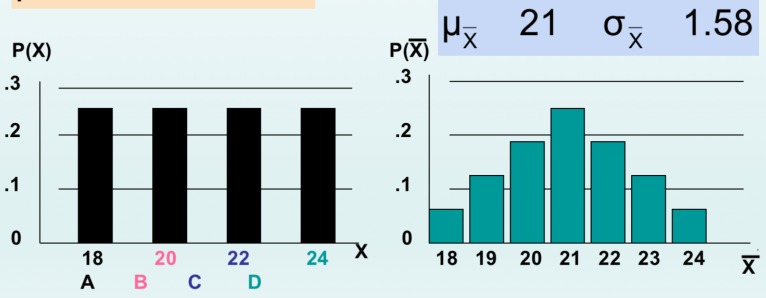


$P(z < -2.5) = 1 - P(z < 2.5) = 1 - F(2.5)$
$= 1 - 0.9938$

Comparing the Population with its Sampling Distribution

Population N = 4: μ 21 σ 2.236

Sample Means Distribution n = 2: $\mu_{\bar{x}}$ 21 $\sigma_{\bar{x}}$ 1.58



Recall Theorem:

If $X_1, X_2, \ldots, X_n$ are independent random variables having normal distributions with means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, respectively, then the random variable

$$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

has a normal distribution with mean

$$\mu_Y = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

and variance

$$\sigma_Y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2.$$

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$E[\bar{X}_1 - \bar{X}_2] = \underbrace{E[\bar{X}_1]}_{\mu_1} - \underbrace{E[\bar{X}_2]}_{\mu_2}$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = var(\bar{X}_1 - \bar{X}_2) = var(\bar{X}_1) + var(\bar{X}_2)$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \Rightarrow \sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Example: Aircraft Engine Life

The effective life of a component used in jet-turbine aircraft engine is a random variable with mean 5000 and SD 40 hours and is close to a normal distribution. The engine manufacturer introduces an improvement into the Manufacturing process for this component that changes the parameters to 5050 and 30. Random samples of size 16 and 25 are selected.

What is the probability that the difference in the two sample means is at least 25 hours?

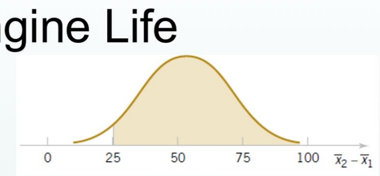


Figure 7-6 The sampling distribution of $\bar{X}_2 - \bar{X}_1$

	Process		
	Old (1)	New (2)	Diff (2-1)
x-bar =	5,000	5,050	50
s =	40	30	
n =	16	25	
Calculations			
s / √n =	10	6	11.7
		z =	-2.14
P(xbar2-xbar1 > 25) = P(Z > z) =			0.9840
			= 1 - NORMSDIST(z)

variance = (standard deviation)^2

$X_1 \sim N(5000, (40)^2) \rightarrow n_1 = 16$

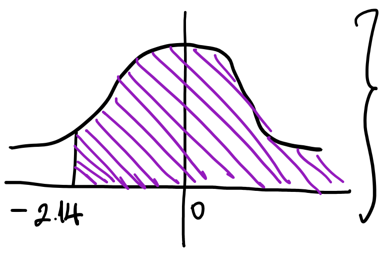
$X_2 \sim N(5050, (30)^2) \rightarrow n_2 = 25$

$P(\bar{X}_2 - \bar{X}_1 > 25) =$

$\bar{X}_2 - \bar{X}_1 \sim N(\mu_2 - \mu_1 = 50, \sigma^2_{\bar{X}_2-\bar{X}_1})$

$\sigma_{\bar{X}_2-\bar{X}_1} = \sqrt{\frac{30^2}{25} + \frac{40^2}{16}} = 11.66 \simeq 11.7$

$P\left(z \geq \frac{25-50}{11.66}\right) = P(z \geq -2.14)$



$P(z \geq -2.14) = P(z < 2.14) = 0.984$