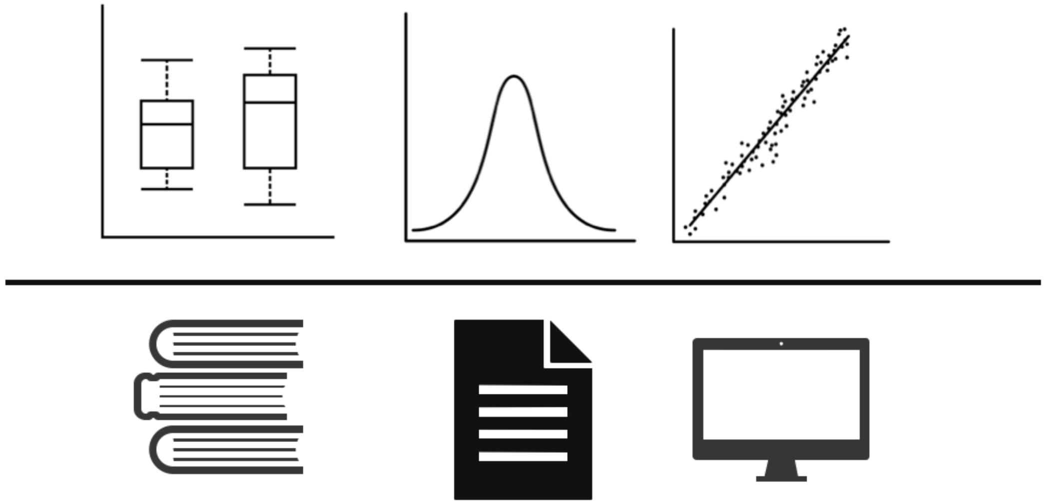
PSYC 2300

Introduction to Statistics



Lecture 08: Midterm Review

Outline for today

- **Mini-Report Assignment**
- **Midterm Review**
  - Review of key computations
  - Review of key concepts
- **Q & A**
  - Any questions about the midterm

Mini-Report Assignment

Reading

The psychology-related article must be from a non-scientific source and must present statistics (descriptive or inferential)

- Psychology Today
- Time
- People

Writing

- 2-pages, double-spaced, 12-point font, 1" margins
- *Summarize* the article and the statistics it reports
- *Describe* how well the author reports these statistics and the methods used to generate them
- *Evaluate* whether a reader who is not scientifically trained would be able to walk away with a clear understanding of the information being presented

Key Computations for the Midterm

- Standard deviation (for samples and population)
- Variance (for samples, populations)
- Cronbach's Alpha
- Correlation
- Z-test statistic
- z-scores

Measures of Variability

Measures of Variability

Measures of variability: describe how scores in a given dataset differ from one another (e.g., the spread or clustering of points)

- Range
- Standard Deviation
- Variance

Measures of Variability

Range: The difference between the lowest and highest values in a dataset

Standard Deviation: The standard (or typical) amount that scores deviate from the mean Variance: The averaged squared deviation from the mean

Variance: The averaged squared deviation from the mean

Sum of Squares

x_i	$x_i - \mu$
1	-1.6
2	-0.6
1	-1.6
4	1.4
3	0.4
3	0.4
6	3.4
1	-1.6
2	-0.6
3	0.4

Step 1: calculate the mean. $\mu = 2.6$

Step 2: subtract each x_i from the mean

Sum of Squares

x	$x_i - \mu$	$(x_i - \mu)^2$
1	-1.6	2.56
2	-0.6	0.36
1	-1.6	2.56
4	1.4	1.96
3	0.4	0.16
3	0.4	0.16
6	3.4	11.56
1	-1.6	2.56
2	-0.6	0.36
3	0.4	0.16

Step 1: calculate the mean. $\mu = 2.6$

Step 2: subtract each x_i from the mean

Step 3: Square the deviations

Step 4: Sum the squared deviations, $\sum(x_i - \mu)^2 = 22.4$

Standard Deviation and Variance

Population Parameter

Standard Deviation

$$\sigma = \sqrt{\frac{\sum(X_i - \mu)^2}{N}}$$

Variance

$$\sigma^2 = \frac{\sum(X_i - \mu)^2}{N}$$

Sample Statistic

Standard Deviation

$$s = \sqrt{\frac{\sum(x_i - \overline{x})^2}{(n-1)}}$$

Variance

$$s^2 = \frac{\sum(x_i - \overline{x})^2}{(n-1)}$$

Cronbach's Alpha

Cronbach's Alpha

Cronbach's Alpha

$$\alpha = (\frac{k}{k-1})(\frac{s_y^2 - \sum s_i^2}{s_y^2})$$

k the number of items

s_y^2 the variance associated with the observed score "total column"

$\sum s_i^2$ the sum of all the variances for each individual item

Cronbach's Alpha

Item_1	Item_2	Item_3
6	6	8
5	5	6
9	8	6
3	2	4
2	3	2
1	1	2
5	4	6

1. I often have worrying thoughts
2. I often feel nervous
3. My heart often beats fast as fears enter in

Cronbach's Alpha

Item_1	Item_2	Item_3	Total
6	6	8	20
5	5	6	16
9	8	6	23
3	2	4	9
2	3	2	7
1	1	2	4
5	4	6	15

Step 1: Sum the values of Item 1, Item 2, and Item 3 *row-wise* to create a "Total" column

Cronbach's Alpha

Item_1	Item_2	Item_3	Total
6	6	8	20
5	5	6	16
9	8	6	23
3	2	4	9
2	3	2	7
1	1	2	4
5	4	6	15

Step 1: Sum the values of Item 1, Item 2, and Item 3 *row-wise* to create a "Total" column

Step 2: Calculate the *sample variance* for each item and 'total'

$$s_i^2 = \frac{\sum(x_i - \overline{x})^2}{(n-1)}$$

$$s_y^2 = \frac{\sum(x_i - \overline{x})^2}{(n-1)}$$

Cronbach's Alpha

Create a table for item 1

Item_1	deviation	squared_deviation
6	1.57	2.47
5	0.57	0.33
9	4.57	20.90
3	-1.43	2.04
2	-2.43	5.90
1	-3.43	11.76
5	0.57	0.33

Step 2a: Calculate the *sample variance* for each item

$$\bar{x} = 4.43$$

$$s_1^2 = \frac{\sum(x_i - \bar{x})^2}{(n-1)}$$

$$s_1^2 = \frac{43.71}{(7-1)} = 7.29$$

Cronbach's Alpha

Create a table for item 2

Item_2	deviation	squared_deviation
6	1.86	3.45
5	0.86	0.73
8	3.86	14.88
2	-2.14	4.59
3	-1.14	1.31
1	-3.14	9.88
4	-0.14	0.02

Step 2b: Calculate the *sample variance* for each item

$$\overline{x} = 4.14$$

$$s_2^2 = \frac{\sum(x_i - \overline{x})^2}{(n-1)}$$

$$s_2^2 = \frac{34.86}{(7-1)} = 5.81$$

Cronbach's Alpha

Create a table for item 3

Item_3	deviation	squared_deviation
8	3.14	9.88
6	1.14	1.31
6	1.14	1.31
4	-0.86	0.73
2	-2.86	8.16
2	-2.86	8.16
6	1.14	1.31

Step 2c: Calculate the *sample variance* for each item

$$\bar{x} = 4.86$$

$$s_3^2 = \frac{\sum(x_i - \bar{x})^2}{(n-1)}$$

$$s_3^2 = \frac{30.86}{(7-1)} = 5.14$$

Cronbach's Alpha

Create a table for the total column

Total	deviation	squared_deviation
20	6.57	43.18
16	2.57	6.61
23	9.57	91.61
9	-4.43	19.61
7	-6.43	41.33
4	-9.43	88.90
15	1.57	2.47

Step 2c: Calculate the *sample variance* for each item

$$\bar{x} = 13.43$$

$$s_y^2 = \frac{\sum(x_i - \bar{x})^2}{(n-1)}$$

$$s_y^2 = \frac{293.71}{(7-1)} = 48.95$$

Cronbach's Alpha

Item_1	Item_2	Item_3	Total
6	6	8	20
5	5	6	16
9	8	6	23
3	2	4	9
2	3	2	7
1	1	2	4
5	4	6	15

Step 3: Prep all of the variance values we just calculated

Item_1	Item_2	Item_3	Total
7.286	5.81	5.143	48.952

$k = 3$ items in our scale

Cronbach's Alpha

Cronbach's Alpha

$$\alpha = (\frac{k}{k-1})(\frac{s_y^2 - \sum s_i^2}{s_y^2})$$

$$\alpha = (\frac{3}{3-1})(\frac{48.95 - 18.24}{48.95})$$

$$\alpha = .941$$

$k = 3$

$\sum s_i^2 = s_1^2 + s_2^2 + s_3^2 = 7.29 + 5.81 + 5.14 = 18.24$

$s_y^2 = 48.95$

Pearson's Correlation

Correlation: Calculation

Pearson's *r*

$$r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

We just need six values:

$\sum X$
$\sum Y$
$\sum XY$

$\sum X^2$
$\sum Y^2$
n

Correlation: Calculation

Set up our table with the x and y values

x	y	x^2	y^2	xy
5	6			
9	11			
10	6			
3	4			
5	6			
7	9			

Correlation: Calculation

Square our x values

x	y	x^2	y^2	xy
5	6	25		
9	11	81		
10	6	100		
3	4	9		
5	6	25		
7	9	49		

Correlation: Calculation

Square our y values

x	y	x^2	y^2	xy
5	6	25	36	
9	11	81	121	
10	6	100	36	
3	4	9	16	
5	6	25	36	
7	9	49	81	

Correlation: Calculation

Multiply x and y together row-wise

x	y	x^2	y^2	xy
5	6	25	36	30
9	11	81	121	99
10	6	100	36	60
3	4	9	16	12
5	6	25	36	30
7	9	49	81	63

Correlation: Calculation

x	y	x^2	y^2	xy
5	6	25	36	30
9	11	81	121	99
10	6	100	36	60
3	4	9	16	12
5	6	25	36	30
7	9	49	81	63

Sum each of the columns

$\sum x$	$\sum y$	$\sum x^2$	$\sum y^2$	$\sum xy$
39	42	289	326	294

Correlation: Calculation

Pearson's *r*

$$r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

$\sum x$	$\sum y$	$\sum x^2$	$\sum y^2$	$\sum xy$
39	42	289	326	294

$$r_{xy} = \frac{(6)(294) - (39)(42)}{\sqrt{[(6)(289) - (39)^2][(6)(326) - (42)^2]}}$$

Correlation: Calculation

$$r_{xy} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

$$r_{xy} = \frac{(6)(294) - (39)(42)}{\sqrt{[(6)(289) - (39)^2][(6)(326) - (42)^2]}}$$

$$r_{xy} = \frac{126}{\sqrt{(213)(192)}} = 0.62306 = 0.62$$

z-test statistic

z-test Statistic

The sample mean we collected

Population mean assuming H_0 is true

z-test statistic

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

Standard error

z-test Statistic

z-test statistic

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}}$$

Standard Error

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$\overline{x}$ = sample mean
$\mu_{\overline{x}}$ = population mean when H_0 is true
$\sigma_{\overline{x}}$ = standard error

σ = population standard deviation
n = sample size

Back to our experiment

In the population

$\mu = 100$

$\sigma = 15$

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$\begin{gathered}101, 122, 132\\ 94, 129, 89\\ 109, 92, 100\\ 125, 103, 91\\ 94, 116, 92\end{gathered}$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{\overline{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$z_{\overline{x}} = \frac{105.9 - 100}{\sigma_{\bar{x}}}$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

$$z_{\overline{x}} = \frac{105.9 - 100}{\sigma_{\bar{x}}}$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\overline{x}} = \frac{15}{\sqrt{15}}$$

$$\sigma_{\overline{x}} = 3.87$$

Back to our experiment

In the population

μ = 100

σ = 15

Our sample

$$n = 15$$

$$\overline{X}_{IQ} = 105.9$$

$$s_{IQ} = 15.10$$

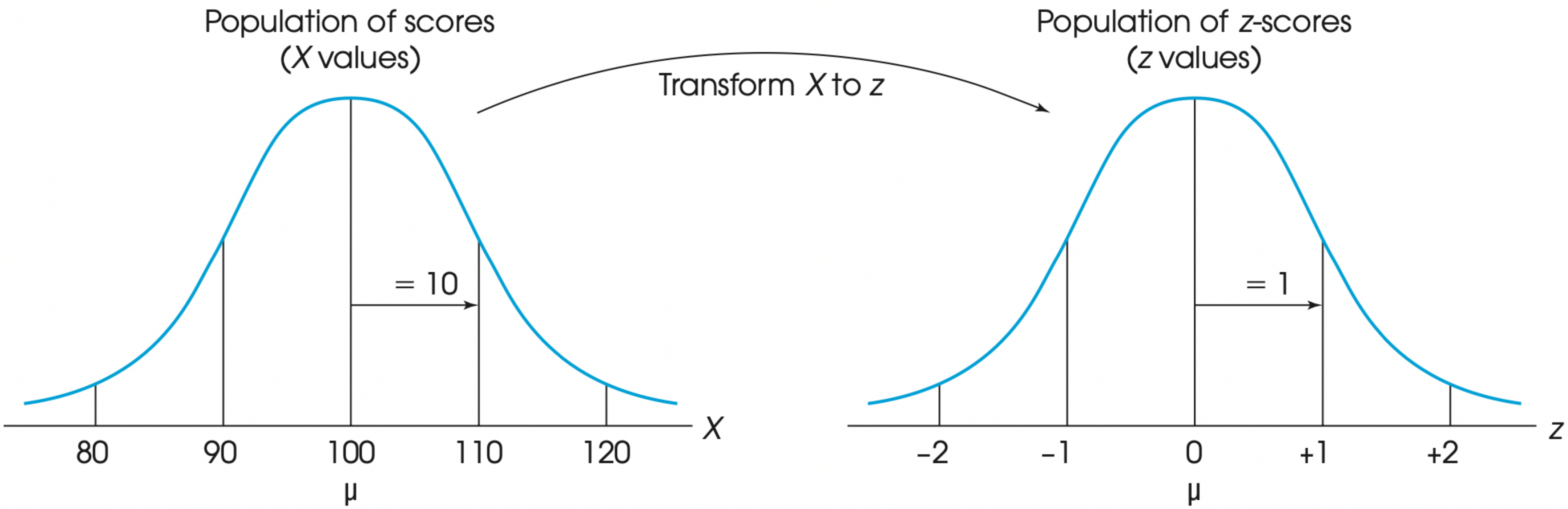
$$z_{\overline{x}} = \frac{105.9 - 100}{3.87}$$

$$z_{\overline{x}} = \frac{5.90}{3.87}$$

$$z_{\overline{x}} = 1.53$$

z-scores

Standardizing Scores



Developing z-scores

Z-score

$$z_i = \frac{x_i - \mu}{\sigma}$$

$$z_i = \frac{deviation_i}{\sigma}$$

The number of standard deviations a particular score deviates from its corresponding mean

z-scoring is useful because a *z*-score of 1 *means the same thing in any distribution*

z-scores

A report in 2010 indicates that Americans between the ages of 8 and 18 spend an average of $\mu = 7.5$ hours per day using some sort of electronic device such as smart phones, computers, or tablets. Assume that the distribution of times is normal with a standard deviation of $\sigma = 2.5$ hours. Aidan checked the 'screen time' info on his phone over the last month, and found that in the last four weeks, he spent $x = 6.3$ hours on the phone per day.

$$\mu = 7.5$$

$$\sigma = 2.5$$

$$x = 6.3$$

$$z_i = \frac{6.3 - 7.5}{2.5} = -0.48$$

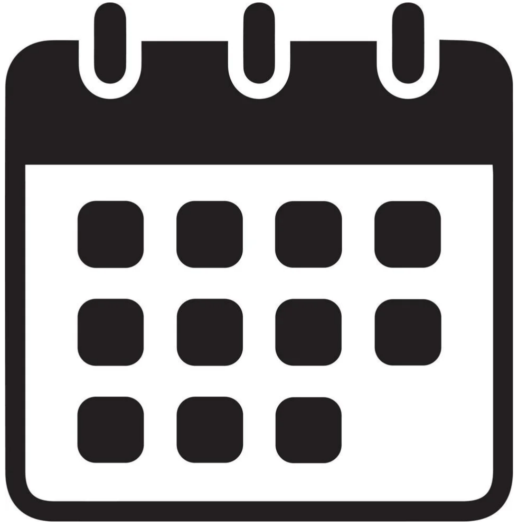
Key concepts for the Midterm

- Mean, median, and mode and their relationship to skewed distributions
- Scales of measurement
- Reliability
- Alpha, *p*-value, critical region
- Type I and Type II errors
- Factors that influence statistical power

Next time

Midterm Exam

- Wednesday 02/02/2022 08:00am MT Sturm Hall 187
  - Bring calculator and pencil
  - Scantron, Formula Sheet, and scratch paper will be provided with the exam



Mini-Report

- Due Wednesday 02/02/2022 11:59pm MT via Canvas
  - Submit `.pdf` or `.docx` file