**01e Review of Basic Statistical Concepts and Introduction to SPSS**

**1. Review of Foundational Concepts**

**Mean and Variance**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Poor | Fair | Good | Very Good | Excellent |
| Overall, how would you rate this instructor? | 1 | 2 | 3 | 4 | 5 |

Ratings from a sample of 5 students:

1, 2, 3, 4, 5

**Mean (M)**

Measure of central tendency; average or typical score

Mean (M) = (Sum of scores) / n

Mean (M) = (1 + 2 + 3 + 4 + 5) / 5

 = 15 / 5

= 3.00

**Variance (s2 or VAR)**

Measure of variation or spread of scores

Variance = (Sum of Squared Deviations) / (degrees of freedom)

Variance = (Sum of Squares) / (degrees of freedom)

Variance = (SS) / (df)

|  |  |  |  |
| --- | --- | --- | --- |
| Scores (X) | Mean (M) | X - M | (X-M)^2 |
| 1 | 3.00 | -2 | 4 |
| 2 | 3.00 | -1 | 1 |
| 3 | 3.00 | 0 | 0 |
| 4 | 3.00 | 1 | 1 |
| 5 | 3.00 | 2 | 4 |
|  |  |  |  |
|  |  | SS = | 10 |

Degrees of Freedom (df) for sample variance = n - 1

Variance = (SS) / (df)

Variance = (10) / (4)

Variance = 2.50

**Standard Deviation (SD or s)**

Measure of score deviation from mean in original units.

SD = $\sqrt{Variance}$

SD = $\sqrt{2.50}$

SD = 1.5811

**Range**

Difference between Maximum and Minimum score. Ignores other scores so not sensitive to variability within max and min scores.

R = Max - Min

R = 5 - 1 = 4

**SPSS**

For practice entering data into SPSS, enter the five student rating scores shown earlier:

1, 2, 3, 4, 5

We can use SPSS to check the values of M and VAR calculated above.

(1) Download and open this SPSS data file:

<http://www.bwgriffin.com/anova/anovadata.sav>

(2) Select the **Variable View** tab then create a new variable called **Scores **

(3) Select **Data View** tab and enter the five ratings in the **Scores** column

****

(4) **Select Analyze → Descriptive Statistics → Frequencies**



(5) Move **Scores** to **Variables** box, then select **Statistics**



(6) Select **Mean** and **Variance**



(7) To obtain results, click **Continue** then **OK**



**Systolic Blood Pressure**

Find M, Md, Mo, SD, VAR, R, and Frequency Distribution for systolic\_bp.

**2. Alpha, Beta, Power, and Others**

**alpha (**$α$**)**

* probability of a Type 1 error;
* normally .05 or .01, values .10 and .001 occasionally used;
* also called significance level

Example:

There’s a 5% chance we will find a relationship between hours studied and achievement in our sample when there isn’t a relationship in the population

Based upon our sample evidence, there’s a 1% chance we will claim our new teaching strategy is more effective when it is not more effective in the population

**beta (**$β$**)**

* probability of a Type 2 error;
* researchers often strive to set this rate .20 or less

Example:

There’s a 20% chance we will claim there is no relationship between hours studied and achievement when there is a relationship in the population

There’s a 10% chance we will claim our new teaching strategy is not more effective when it is more effective in the population

**power (1-**$β$**)**

* probability of identifying an effect if one exists;
* probability of rejecting a false null hypothesis;
* complement of beta (**1-**$β$) so it is the probability of not committing a Type 2 error

Example:

There’s an 80% chance we will detect a relationship between hours studied and achievement in our sample if there is actually a relationship in the population

With this small sample there is only a 40% chance we will correctly claim our new teaching strategy is more effective when it is more effective

**n** is study sample size

Example:

In our pregnancy study we tested 15 doctors to determine if they could correctly identify pregnancy status of patients

To test our new teaching strategy, we sampled 108,637 students; 46,714 students received the new teaching strategy and 61,923 received current standard instructional approaches

**effect size**

* effect size denotes the magnitude of difference if comparing means or magnitude of relationship between variables
* may be unstandardized or standardized
	+ unstandardized = females earn $2,354 more than males
	+ standardized = females earn 0.43 SD more than males

Example:

Students exposed to the new teaching strategy scored, on average, 0.12 points (unstandardized difference) higher on unit tests than students exposed to traditional instruction

Mean New Strategy = 83.35

Mean Traditional Teaching = 83.23

Mean Difference = 00.12

With a SD = 7, the corresponding standardized effect size is

d = .017

The relationship between academic self-efficacy and test anxiety is b = -.16 (unstandardized regression coefficient) and r = -.45 (standardized effect size; Pearson correlation)

**significant result**

* simply means Ho rejected;
* does not mean results are important;
* larger samples are likely to lead to significant results even when effect sizes are trivial

Example:

There is a statistically significant difference in student achievement between those exposed to the new teaching strategy and those exposed to traditional instruction (mean achievement scores: 83.35 vs. 83.23, SD = 7.00, d = .017).

(Note: To have an 80% chance of declaring this small difference “significant” at the .05 level requires a sample of 108,637 students)

**insignificant result**

* simply means Ho was not rejected;
* does not mean results are unimportant

Example:

There is not a statistically significant difference in red blood cell destruction counts between those taking Soliris (nonfictional drug) and those taking the rival experimental OxygenCell (fictional drug).

For this study why would failure to reject the null be important?

|  |
| --- |
| “Alexion Pharmaceutical’s Soliris, at $409,500 a year, is the world’s single most expensive drug. This monoclonal antibody drug treats a rare disorder in which the immune system destroys red blood cells at night. The disorder, paroxysymal nocturnal hemoglobinuria (PNH), hits 8,000 Americans. Last year Soliris sales were $295 million.”Source: www.forbes.com/sites/matthewherper/2012/09/05/how-a-440000-drug-is-turning-alexion-into-biotechs-new-innovation-powerhouse |

What if OxygenCell costs only $120 per year? Then failure to find a difference (i.e., failure to reject the null) in red blood cell destruction counts between OxygenCell and Soliris means equivalent effectiveness at a cost savings.

**p-value**

Assuming Ho is true, the p-value is the probability of randoming obtaining sample results, for a sample of size n, similar to or more extreme from those data sampled

Example:

I began taking blood pressure (BP) medication a few years ago. I measured my heart rate (beats per minute) each of 20 mornings before I began taking BP medication and then again 20 mornings after taking BP medication. Results are presented below.

|  |  |  |
| --- | --- | --- |
|  | Before Medication | After Medication |
| Mean | 55.357 | 49.714 |
| SD | 3.300 | 5.00 |
| n | 20 | 20 |
|  |  |  |
| t-ratio | 3.526 |  |
| df | 26 |  |
| p-value | 0.002 |  |

Interpretation of p-value 0.002:

If in reality there is no difference in my heart rate before and after taking BP medication (i.e., assume Ho is true), then the probability of obtaining a sample of 40 heart rate readings at random with a difference of 55.357-49.714 = 5.64 beats per minute is 0.002 (about 2 times out of 1,000 sampling attempts).

**3. Errors in Hypothesis Testing; Type 1 vs. Type 2**

**Type 1 error**

* + incorrectly rejecting a true null hypothesis;
	+ claiming there is an effect based upon sample results when there is not an effect in the population;
	+ a false positive

Example

Using sample data one concludes a new teaching strategy produces higher achievement than traditional practices when, in fact, it does not show this difference in the population

**Type 2 error**

* failure to reject a false null hypothesis;
* failure to identify an effect in the sample when there is an effect in the population;
* a false negative

Example

Based upon a sample one claims a new teaching strategy does not produce higher achievement than traditional practices when, in fact, it does in the population

**Type 1 and 2 Errors Illustrated**

 Assume the following null hypothesis:

Ho: patient is not pregnant.

***Question 1***

Which of the following diagnoses would be Type 1 and Type 2 errors?



Image source

<http://effectsizefaq.com/2010/05/31/i-always-get-confused-about-type-i-and-ii-errors-can-you-show-me-something-to-help-me-remember-the-difference/>

Display idea

Ellis, Paul D. (2010). The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results.