**Stats Medic Ultimate Interpretations Guide**

**CED Unit 1: Exploring One-Variable Data**

**Standard Deviation**: The **context** typically varies by SD from the mean of mean.

Example: *The* ***height of power forwards in the NBA*** *typically varies by 1.52 inches from the*

 *mean of 80.1 inches.*

**Percentile**: percentile % of **context** are less than or equal to value.

 Example: *75% of* ***high school student SAT scores*** *are less than or equal to 1200.*

**z-score:** Specific value with **context** is z-score standard deviations above/below the mean.

 Example: *A* ***quiz score*** *of 71 is 1.43 standard deviations below the mean. (z = -1.43)*

**Describe a distribution:** Be sure to address shape, center, variability, and outliers (in **context**).

 Example: *The* ***distribution of student height*** *is unimodal and roughly symmetric. The mean height is 65.3 inches with a standard deviation of 8.2 inches. There is a potential upper outlier at 79 inches and a gap between 60 and 62 inches.*

**CED Unit 2: Exploring Two-Variable Data**

**Correlation** $\left(r\right)$**:** The linear association between **x-context** and **y-context** is weak/moderate/strong (strength) and positive/negative (direction).

 Example: *The linear association between* ***student absences*** *and* ***final grades*** *is fairly strong and negative. (r =* −*0.93)*

**Residual:** The actual **y-context** was residual above/below the predicted value when **x-context** = #.

 Example: *The actual* ***heart rate*** *was 4.5 beats per minute above the number predicted when* ***Matt ran*** *for 5 minutes.*

**y-intercept:** The predicted **y-context** when x = 0 **context** is y-intercept.

 Example: *The predicted* ***time to checkout at the grocery store*** *when there are 0* ***customers in line*** *is 72.95 seconds.*

**Slope**: The predicted **y-context** increases/decreases by slope for each additional **x-context**.

 Example: *The predicted* ***heart rate*** *increases by 4.3 beats per minute for each additional* ***minute jogged****.*

**Standard Deviation of Residuals (*s*):** The actual **y-context** is typically about s away from the value predicted by the LSRL.

 Example: *The actual* ***SAT score*** *is typically about 14.3 points away from the value predicted by the LSRL.*

**Coefficient of Determination** $\left(r^{2}\right)$: About $r^{2}$% of the variation in **y-context** can be explained by the linear relationship with **x-context**.

Example: *About 87.3% of variation in* ***electricity production*** *is explained by the linear*

 *relationship with* ***wind speed****.*

**Describe the relationship**: Be sure to address strength, direction, form and unusual features (in **context**).

 Example: *The scatterplot reveals a moderately strong, positive, linear association between the* ***weight and length of rattlesnakes****. The point at (24.1, 35,7) is a potential outlier.*

**CED Unit 4: Probability, Random Variables and Probability Distributions**

**Probability** $P(A)$: After many many **context**, the proportion of times that **context A** will occur is about P(A).

 Example: P(heads) = 0.5.

 *After many many* ***coin flips****, the proportion of times that* ***heads*** *will occur is about 0.5.*

**Conditional Probability** $P(A|B$**)**: Given **context B**, there is a P(A|B) probability of **context A**.

 Example: P(red car | pulled over) = 0.48.

 *Given that* ***a car is pulled over****, there is a 0.48 probability of* ***the car being red****.*

**Expected Value (Mean,** $μ$**)**: If the random process of **context** is repeated for a very large number of times, the average number of **x-context** we can expect is expected value. (decimals OK).

 Example: *If the random process of* ***asking a student how many movies they watched this week*** *is repeated for a very large number of times, the average number of* ***movies*** *we can expect is 3.23 movies.*

**Binomial Mean** $\left(μ\_{X}\right)$: After many, many trials the average # of **success context** out of n is $μ\_{X}$.

 Example: *After many, many trials the average # of* ***property crimes that go unsolved*** *out of*

 *100 is 80.*

**Binomial Standard Deviation** $\left(σ\_{X}\right)$: The number of **success context** out of n typically varies by $σ\_{X}$ from the mean of $μ\_{X}$.

 Example: *The number of* ***property crimes that go unsolved*** *out of 100 typically varies by 1.6*

 *crimes from the mean of 80 crimes.*

**CED Unit 5: Sampling Distributions**

**Standard Deviation of Sample Proportions (**$σ\_{\hat{p}}$**)**: The sample proportion of **success context** typically varies by $σ\_{\hat{p}}$ from the true proportion of $p$.

 Example: *The sample proportion of* ***students that did their AP Stats homework last night***

 *typically varies by 0.12 from the true proportion of 0.73.*

**Standard Deviation of Sample Means (**$σ\_{\overbar{x}}$**)**: The sample mean amount of **x-context** typically varies by $σ\_{\overbar{x}}$ from the true mean of $μ\_{X}$.

 Example: *The sample mean amount of* ***defective parts*** *typically varies by 5.6 parts from the true mean of 23.2 parts.*

**CED Unit 6, 7, 8 & 9: Inference for Proportions, Means, and Slope**

**Confidence Interval (A, B)**: We are % confident that the interval from A to B captures the true **parameter context**.

 Example: *We are 95% confident that the interval from 0.23 to 0.27 captures the true* ***proportion of flowers that will be red after cross-fertilizing red and white****.*

**Confidence Level**: If we take many, many samples of the same size and calculate a confidence interval for each, about confidence level % of them will capture the true **parameter in context**

 Example: *If we take many, many samples of size 20 and calculate a confidence interval for each, about 90% of them will capture the true* ***mean weight of a soda case****.*

***p*-value**: Assuming $H\_{0}$ **in context** $(H\_{0})$, there is a *p*-value probability of getting the observed result or less/greater/more extreme, purely by chance.

Example: *Assuming the* ***mean body temperature is 98.6* °*F*** *(*$H\_{0}: μ $*= 98.6), there is a*

 *0.023 probability of getting a sample mean of 97.9* °*F or less, purely by*

 *chance.*

**Conclusion for a Significance Test**: Because *p*-value *p*-value < / > $α$ we reject / fail to reject H0. We do / do not have convincing evidence for $H\_{a }$ **in context**.

 Example: *Because the* p*-value 0.023 < 0.05, we reject H0. We do have convincing evidence that the* ***mean body temperature is less than 98.6* °*F*** $\left(H\_{a}: μ<98.6\right)$.

**Type 1 Error**: The $H\_{0}$ **context** is true, but we find convincing evidence for $H\_{a}$ **context**.

 Example: *The* ***mean body temperature is actually 98.6* °*F****, but we find convincing*

 *evidence* ***the mean body temperature is less than 98.6* °*F****.*

**Type II Error**: The $H\_{a}$ **context** is true, but we don’t find convincing evidence for $H\_{a}$ **context**.

 Example: *The* ***mean body temperature is actually less than 98.6* °*F****, but we don’t find*

 *convincing evidence that* ***the mean body temperature is less than 98.6* °*F.***

**Power**: If $H\_{a}$ **context is true at a specific value** there is a power probability the significance test will correctly reject $H\_{0}$.

 Example: *If the* ***true mean body temperature is 97.5* °*F****, there is a 0.73 probability the significance test will correctly reject* $H\_{0}: μ $*= 98.6*

**Standard Error of the Slope** $(SE\_{b})$: The slope of the sample LSRL for **x-context** and **y-context** typically varies from the slope of the population LSRL by about $SE\_{b}$.

 Example: *The slope of the sample LSRL for* ***absences*** *and* ***final grades*** *typically varies from the slope of the population LSRL by about 1.2 points/absence.*