**Key Concepts Psych 3020**

***This is a very brief review of the major concepts in this course. It is not an all-encompassing review of course or final exam material.***

1. Sample Versus Population

* Instead of collecting data on every individual in a population (not possible) we collect data on a sample of the population, which represents a portion of the total population.
* The sample mean in an unbiased estimator of the population mean (when random sampling is used).
* A statistic is a number that describes some characteristic of a sample.
* A parameter is a number that describes some characteristic of a population.

| **Basis for Comparison** | **Sample Mean** | **Population Mean** |
| --- | --- | --- |
| Meaning | Sample mean is the arithmetic mean of random sample values drawn from the population.  | Population mean represents the actual mean of the whole population. |
| Symbol | x̄ (pronounced as x bar) | μ (Greek term mu) |
| Calculation | Easy | Difficult |
| Accuracy | Low | High |
| Standard deviation | When calculated using sample mean, is denoted by (s). | When calculated using population mean, is denoted by (σ). |

2. Types of Variables

* Discrete variables: They are counting numbers like 0,1,2,3
* Continuous variables: They can be any type of value including fractions and decimals.

3. Hypothesis Testing

* Use a specific testable statement with direction (i.e., Therapy will significantly reduce depression, *NOT*, therapy will predict levels of depression).
* If there is sufficient evidence to reject the null hypothesis, we conclude the alternative hypothesis is true.
* If there is not sufficient evidence to reject the null hypothesis, we conclude the null hypothesis might be true.
* You can base your conclusion on the Level of significance (α) and the p-value. Remember the level of significance can be determined from the level of confidence (level of significance (α) = 1- level of confidence (c))
	+ **Typically α = .05**
* If p ≤ α, reject the null hypothesis (Ha is true)
* If p > α, fail to reject the null hypothesis (H₀ is true)



4. Error in Hypothesis Testing

|  |  |
| --- | --- |
|  | **REALITY** |
| **Decision** | **Null hypothesis is true** | **Null hypothesis is false** |
| **Reject null hypothesis** | Type I Error | Correct Decision |
| **Don’t reject null hypothesis** | Correct Decision | Type II Error |

5. Probability

* OR EVENTS-**-** An outcome is in the event A OR B if the outcome is in A or is in B or is in both A and B.
* AND EVENTS--An outcome is in the event A AND B if the outcome is in both A and B at the same time.
* CONDITIONAL PROBABILITY--The conditional probability of A given B is written P(A|B). P(A|B) is the probability that event A will occur given that the event B has already occurred.

6. Chi Square Test

* Non-parametric test.
* Chi-square distributions df based on number of cells not *n.*
* Goodness of fit of distributions.
* Test of independence.
* Test of homogeneity.
* 5 per cell minimum.

7. *t-*distribution

* Hypothesis testing for population means σ is unknown

|  |  |
| --- | --- |
| **Letter** | **Meaning** |
| µ | Population mean is the mean that is used in the hypotheses |
| $$\overbar{x}$$ | Sample mean ($\overbar{x}$) is the mean of the sample |
| $$\overbar{X}$$ | Random variable. Different from sample mean. Represents the average for what is being investigated. |
| n | Sample size |
| Sx | sample standard deviation |
| α | Level of significance  |
| c | Level of confidence |
| Distribution | Student-t distribution written as $t\_{n-1}$ |

* One-sample *t*-test compares a sample mean against a known value in the population.
* Independent samples *t*-test compares means between two groups.
* Paired samples *t*-test compares differences in means across different measurement occasions (same sample).

8. Normal Distribution

* A normal distribution naturally arises from multiply-determined events.
* Hypothesis testing, population mean is known = *Z-*Test.
* **The z-score tells you how many standard deviations the value x is above (to the right of) or below (to the left of) the mean, μ.**
* *Z ~ N (mean from hypotheses,* $\frac{population standard devation}{\sqrt{sample size}}$*)*



9. Linear regression and Correlation

* Independent variable is the predictor or causal variable
* Dependent variable *depends* on independent variable; it is the outcome variable
* y= mx+b
	+ b = slope
	+ a = y-intercept
	+ x = IV
	+ y = DV
* r² = explains variation of y explained by x.
* r = correlation coefficient
	+ Strength of the relationship
	+ Cannot be used on its own to determine causality

10*. F*-distribution and Analysis of Variance

* Right tailed test.
* Compare more than 2 groups.
* Ho: $µ\_{1}$ = $µ\_{2}$ = $µ\_{3}$ = ……$µ\_{n}$ (which means all the means equal each other).
* Ha: At least two of the group means are not equal.
* Factorial design.
* Variance within groups and variance between groups.
* Number of groups = *k*
* Total sample size across all groups = *n*
* Degrees of freedom (numerator) = *k* – 1
* Degrees of freedom (denominator) = *n – k*

11. Reliability and Validity

* Reliability refers to a statistical property of the scale. Reliability indexes the percentage of the construct that is captured by the items.
* Reliability is expressed as a proportion, so it ranges from 0 to 1.
* “high” α ≥ .85
* “good” α ≥ .80
* “reliable” α between .7 and .8
* “adequate” α between .6 and .7
* “poor” α ≤ .6
* Validity
	+ Face validity: Do the items look like they are measuring what they purport to measure? For example, if we are asking about participants’ meaning in life, do the items appear to be related to this construct?
	+ Content validity: Do the items cover the entire range of the construct? Psychological constructs can be very specific or quite broad.
	+ Criterion/Predictive validity: Does the psychological scale correlate well with other measures that it should logically correlate with?



12. Discrete Random Variables

* Probability distribution functions.
* Expected values.
	+ μ = ∑ xP(x).
* Binomial distributions.
	+ There are a fixed number of trials, n.
	+ There are only two possible outcomes, called "success" and, "failure," for each trial. The letter p denotes the probability of a success on one trial, and q denotes the probability of a failure on one trial.
	+ The n trials are independent and are repeated using identical conditions.
* Geometric distributions.
	+ There are one or more Bernoulli trials with all failures except the last one, which is a success.
	+ In theory, the number of trials could go on forever. There must be at least one trial.
	+ The probability, p, of a success and the probability, q, of a failure do not change from trial to trial.
* Hypergeometric distributions.
	+ A fixed number of trials.
	+ The probability of success is not the same from trial to trial.
* Poisson distributions.
	+ A fixed number of trials.
	+ The probability of success is not the same from trial to trial.

13. Continuous Random Variables

* Continuous probability functions
* Uniform Distribution
	+ If X has a uniform distribution where a < x < b or a ≤ x ≤ b, then X takes on values between a and b (may include a and b). All values x are equally likely. We write X ∼ U(a, b).
* Exponential Distribution
	+ The exponential distribution has the **memoryless property,** which says that future probabilities do not depend on any past information.