The Normal Distribution and Standard Scores

Why is the normal distribution important?	1. Many naturally occurring data (e.g., height, weight, etc,) have many distributions which are approximately normal. 2. Many statistical tests covered later use normal distributions. 3. Many sampling distributions approximate a normal distribution with large sample sizes.
Properties of a normal distribution	- Unimodal - Mean is middle most score - Equal on each side -Two injection points occurring at (x μ+1σ & μ–1σ)
Area under the normal distribution	Calculated in percentages, the total area under the curve = 100%. Broken up into 8 sections. (0.13, 2.15, 13.9, 34.13, 34.13,(mean (No
Area under the normal curve it's based on	The number of standard deviations from the mean is constant for all normal distributions.
For any score…	If we know how many standard deviations it is away from the mean
How do we calculate?	$z = (X-\mu)/\sigma$

Z Scores

What is a standard (or z) Score?	z score is a *transformed* score that designates how many standard deviation units the corresponding raw score is above or below the mean.
What are the properties of z scores?	1. Mean=0 (μ_z=0) 2. Standard deviation=1 (σ_z=1) 3. Shape of z score distribution is the **SAME** as shape of raw score distribution -> The relative positions of the scores in the distribution do not change either
Column A	Shows the z score
Column B	Area between mean and z
Column C	Area beyond z
Column B and C will always add up to...	0.5000
Area under the normal curve based on the number of standard deviations from the mean is...	**constant for all normal distributions**
The scores we calculate are also called	- z score - normal scores - standardized scores*
Converting z scores will...	Standardize any distribution without regard to the original mean or SD
Once it is standardized it will...	Always have a mean of 0 and a SD of 1 which allows for comparison across different distributions

Probability

What are the two types of questions in inferential statistics?	1) Hypothesis testing 2) Parameter estimation

Probability (cont)

Hypothesis testing	We have a hypothesis about a certain population and we wish to test it using a sample drawn from that populations
Parameter estimation	We wish to know the magnitude of a population characteristic, so we test a sample (e.g., how much salary do students who graduate with a psych degree make in Canada?)
The goal is to...	Infer something about the population based on the info from a sample, thereforethis sample has to be representative of the population and it must be **a random sample.**
Random sample	A sample selected from the population that satisfies the following two condition 1) Each possible sample has an equal chance of being selected 2) Each member of the population has an equal chance of being selected into the sample.
Why do we need random samples?	1) If we wish to generalize to the population, the sample must be representative of the population. 2) The laws of probability cannot be used if the sample isn't random
Probability	1) Cannot be negative (between 0-1) - Probability = 0 (event is certain not to occur) - Probability = 1 (event is certain to occur 2)Usually expressed as a decimal number but can be written as a fraction (keep 4 decimal places)
Probability can be calculated in two ways...	1) a priori probability - **deduced from reason** (i.e., theoretically based), without observations 2) A posteriori probability - Calculated **based on the actual observations** (i.e., empirically based)
A priori	From before
A posteriori	After the fact

A priori probability

A priori probability	Based on reason without actual observations
P(A) =	Number of events classifiable as "A"/ Total number of possible events
What is the a priori probability of flipping a coin and getting a "head"	p(A) = 0.5

A posteriori probability

A posteriori probabiity	Based on the actual observations

A posteriori probability (cont)

P(A)	Number of times "A" has actually occurred/ Total number of occurrences
If we actually flipped a coin 50 times, and got a head 30 times, what is the a posteriori probability of getting a "head"	p(A) = 0.60

Multiplication rule for probability

Multiplication rule	Concerned with determining the probability of **joint or successive occurrence** of several events
Multiplication rule example: There are two events (event a , event B) We can ask...	1) What is the probability of both A and B happening together 2) What is the probability of A happening first and B happening second?
P(A)	Probability of A
P(B|A)	Probability of B, given that A has occurred
P(A and B)	P(A)p(B|A)
Independent events	Two events are independent if the occurrene of one event has no effect on the probability of occurrence of the other event Note:sampling with replacement results in INDEPENDENT EVENTS (p(A and B) = p(A)p(B)
Example question: There are two dice. What is the probability of getting a "3" on the 1st die and a "4" on the 2nd die in one roll?	Event A: "3" on the 1st die -p("3" on the 1st die) = 1/6 Event B: "4" on the 2nd die -p("4" on the 2nd die|"3" on the 1st die) = 1/6 {{nl}} (1/6)(1/6) = 0.0278
Dependent events	The two events are dependent if the occurrence of one event (e.g., A) has an effect on the probability of occurrence of the other event (e.g., B). Note: Sampleing WITHOUT replacement results in DEPENDENT EVENTS p(A and B) = p(A)p(B|A)

Addition for probability

Mutually exclusive events	Two events are mutually exclusive when the occurrence of one *precludes* the occurrence of the other. **Two events that CANNOT occur together** p(A and B) = 0
Addition rule for probability	Concerned with determining the probability of occurrence of **any one** of several possible events - Probability of A **or** B
p(A or B) =	p(A) +p(B) - p(A and B)

Addition for probability (cont)

Example: What is the probability that you will draw a king or a diamond on the first card from the deck?	Event A: King on the 1st card - p(king) = 4/52 Event B: Diamond on the 1st card p (diamond) = 13/52 = (4/52) + (13/52) - (1/52) = 16/52 = **0.3077**
Exhaustive sets of events	A set of events is exhaustive if the set includes all of the possible events (rolling a die, the set of events of getting a 1, 2, 3, 4, 5, or 6 is exhaustive; flipping a coin, the set of events of getting a head or tail is exhaustive)
If a set of events (A, B, C ...) are exhaustive and mutually exclusive	p(A) + p(B) + p(C) + ... = 1
Example (M(*)&A(+)): If you have a regular deck of playing cards, what is the probability that **at least one of the next three cards** will be red (w/o replacement)?	p(at least 1 out of 3 red) = 1-p(all black) =1-(26/52)(25/51)(24/50) =1-0.117647 =0.8824

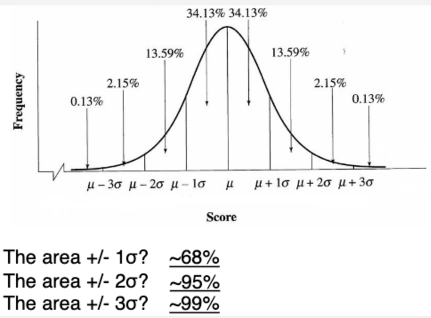
Hypothesis Testing

Why can't we just look at the data?	The varaibility in data, it's very hard to "see" the difference between groups or conditions (could have happened due to chance). This is why we need to use inferential stats to test hypotheses, to determine whether there's a real difference between groups or conditions that is due to IV (or subject variable).
Free throw distractions in Basketball	Do free throw distractions influence the player's ability to successfully make free throws?
Example hypotheses	- Fan distractions affects free throw accuracy (H_1) - Fan distractions does not affect free throw accuracy (H_0) -Free throws are more difficult to make with distractions (H_1) -Free throws are not more difficult ot make with distractions(H_0) - Free throws are easier to make with distractions (H_1) - Free throws ar enot easier to make with distractions (H_0)

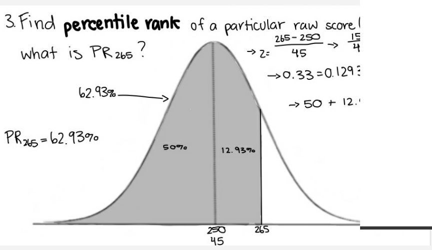
Hypothesis Testing (cont)

Null hypothesis	-hypothesies no effect - No dfiference bwtween groups No difference between conditions no relationship NO DIFFERENCE - NO EFFECT
Alternative hypothesis	- Hypothesizes that ther will be difference between groups / conditions and hat this dfference is due to the indepe-ndent variable/ subject variable
H0 and H1 must be...	mutually exclusive and exhaustive
Decision rule	- there must be criteria by which we will decide3 if the independent variable really did have an effect (we can use probability)
IF the proability is low	We will reject H0 and accept H1
If the probabiliyt is not that low	We will not reject H0 a
Threashold	a (alpha) 0.05 or for more precision 0.01
Type 1 error	Decide to reject eh null hypothesis but the null is actually true
Type 2 error	Decided to keep the null hypothesis but it actually is'nt true.

Breakdown of Normal Distribution Curve

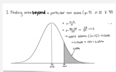


Find percentile rank of a particular raw score

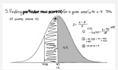


Find actual # of cases below a particular z score

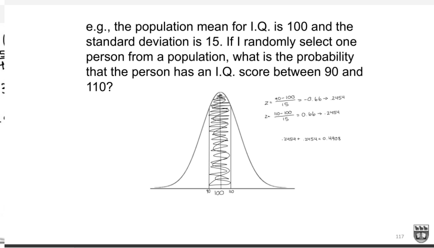
Finding area beyond a particular raw score



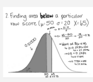
...g particular raw scores of ...n area



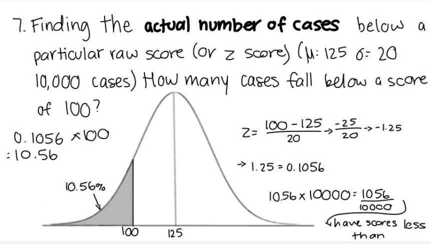
P of normally distributed cont. var. E.g. 2

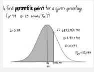


Finding area below a particular raw score

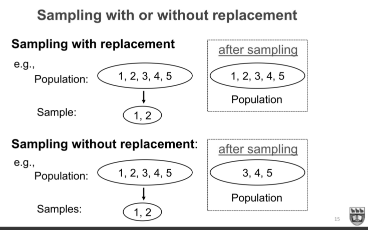


Find percentile point for a given percentage

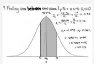




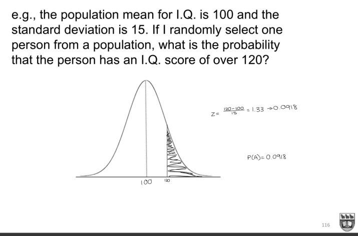
Sampling with or without replacement



Finding area between two raw scores



P of normally distributed cont. var. E.g. 1



p(A) = Area under the curve corresponding ot A / Total area under the curve