0.0.1 Significance Level (α)

- In hypothesis testing, the significance level α is the threshold used for rejecting the null hypothesis.
- The significance level is used in hypothesis testing as follows: First, the difference between the result (i.e. observed statistic or point estimate) of the experiment and the **null value**
- The null value is the expected value of this statistic, assuming that the null hypothesis is true), is determined.(i.e.we denote this ***Observed - Null***).
- Then, assuming the null hypothesis is true, the probability of a difference that large or larger is computed .
- Finally, this probability is compared to the significance level.
- If the probability is less than or equal to the significance level, then the null hypothesis is rejected and the outcome is said to be statistically significant.

0.0.2 Significance (Die Throw Example)

- Suppose that the outcome of the die throw experiment was a sum of 401. In previous lectures, a simulation study found that only in approximately 1.75% of cases would a fair die yield this result.
- However, in the case of a crooked die (i.e. one that favours high numbers) this result would not be unusual.
- A reasonable interpretation of this experiment is that the die is crooked, but importantly the experiment doesn't prove it one way or the other.
- We will discuss the costs of making a wrong decision later (Type I and Type II errors).

0.0.3 Significance Level

- Traditionally, experimenters have used either the 0.05 level (sometimes called the 5% level) or the 0.01 level (1% level), although the choice of levels is largely subjective.
- The lower the significance level, the more the data must diverge from the null hypothesis to be significant.
- Therefore, the 0.01 level is more conservative than the 0.05 level.
- The Greek letter alpha (α) is sometimes used to indicate the significance level.
- We will use a significance level of $\alpha = 0.05$ only in this module. You may assume this level unless clearly stated otherwise

0.0.4 Significance Level

- The significance level of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis H_0, if it is in fact true.
- Equivalently, the significance level (denoted by α) is the probability that the test statistics will fall into the ***critical region***, when the null hypothesis is actually true. (We will discuss the critical region shortly).
- Common choices for α are 0.05 and 0.01

0.1 Significance Level

- In everyday spoken English, "*significant*" means important, while in Statistics "*significant*" means probably true (not due to chance). A research finding may be true without being important.
- When statisticians say a result is "highly significant" they mean it is very probably true.
- 95% is commonly used as the accepted level in statistical college classes. In Medical research, it is likely to be much higher, e.g. 99.9%.
- The significance level of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis H_0, if it is in fact true.
- It is the probability of a ***type I error*** and is set by the investigator in relation to the consequences of such an error. That is, we want to make the significance level as small as possible in order to protect the null hypothesis and to prevent, as far as possible, the investigator from inadvertently making false claims.
- The significance level is usually denoted by α (alpha):

$$\text{Significance Level} = \text{P(type I error)} = \alpha$$

- Usually, the significance level is chosen to be 0.05 (or equivalently, 5%).

0.2 What is Statistical Inference?

- Statistical inference is about inferring from the data about parameters that describe an assumed model for the data.
- Solution: In statistics, a model for the mechanism that has produced data is as- sumed. The model is characterized by some parameters that are unknown. Having data, statistics tries to infer from them some information about these parameters.

0.3 What is Statistical Inference?

- Hyptothesis testing
- Confidence Intervals
- Sample size estimation.

0.4 Statistical significance

- Statistical significance is a mathematical tool used to determine whether the outcome of an experiment is the result of a relationship between specific factors or due to chance.
- Statistical significance is commonly used in the medical field to test drugs and vaccines and to determine causal factors of disease. Statistical significance is also used in the fields of psychology, environmental biology, and any other discipline that conducts research through experimentation.
- Statistics are the mathematical calculations of numeric sets or populations that are manipulated to produce a probability of the occurrence of an event. Statistics use a numeric sample and apply that number to an entire population.
- For the sake of example, we might say that 80% of all Americans drive a car. It would be difficult to question every American about whether or not they drive a car, so a random number of people would be questioned and then the data would be statistically analyzed and generalized to account for everyone.
- In a scientific study, a hypothesis is proposed, then data is collected and analyzed. The statistical analysis of the data will produce a number that is statistically significant if it falls below 5%, which is called the confidence level. In other words, if the likelihood of an event is statistically significant, the researcher can be 95% confident that the result did not happen by chance.
- Sometimes, when the statistical significance of an experiment is very important, such as the safety of a drug meant for humans, the statistical significance must fall below 3%. In this case, a researcher could be 97% sure that a particular drug is safe for human use. This number can be lowered or raised to accommodate the importance and desired certainty of the result being correct.
- Statistical significance is used to reject or accept what is called the null hypothesis. A hypothesis is an explanation that a researcher is trying to prove. The null hypothesis holds that the factors a researcher is looking at have no effect on differences in the data.
- Statistical significance is usually written, for example, $t = .02, p < .05$. Here, ”t” stands for the statistic test score and ”p¡.05” means that the probability of an event occurring by chance is less than 5%. These numbers would cause the null hypothesis to be rejected, therefore affirming that the alternative hypothesis is true.
- Here is an example of a psychological hypothesis using statistical significance: It is hypothesized that baby girls smile more than baby boys. In order to test this hypothesis, a researcher would observe a certain number of baby girls and boys and count how many times they smile. At the end of the observation, the numbers of smiles would be statistically analyzed.
- Every experiment comes with a certain degree of error. It is possible that on the day of observation all the boys were abnormally grumpy.
- The statistical significance found by the analysis of the data would rule out this possibility by 95% if t=.03. In this case, the null hypothesis that baby girls do not smile more than baby boys would be rejected, and with 95% certainty, the researcher could say that girls smile more than boys.

Statistical Inference : Confidence Intervals

- Confidence intervals allow us to use sample data to estimate a parameter value, such as a population mean.
- A confidence interval is a range of values for which we can be confident (at a specific level) that parameter value (such as the population mean) lies within.
- A confidence level will have a specified level of confidence, commonly 95%.
- The 95% confidence interval is a range of values which contains the parameter value of interest with a probability of 0.95.
- We can expected that a 95% confidence interval will not contain the parameter value of interest with a probability of 0.05.

Statistical Inference : Confidence Intervals

- It is natural to interpret a 95% confidence interval on the mean as an interval with a 0.95 probability of containing the population mean.
- However, the proper interpretation is not that simple.
- Consider the case in which 1,000 studies estimating the value of μ in a certain population all resulted in estimates between 30 and 40.
- Suppose one more study was conducted and the 95% confidence interval on μ was reported to be $40 \leq \mu \leq 50$ (based on that one study).
- The probability that μ is between 40 and 50 is very low, the reported confidence interval not withstanding.

Standard Error

- The standard error measures the dispersion of the sampling distribution.
- For each type of point estimate, there is a corresponding standard error.
- A full list of standard error formulae will be attached in your examination paper.
- The standard error for a mean is
$$S.E(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$
However, we often do not know the value for σ. For practical purposes, we use the sample standard deviation s as an estimate for σ instead.
$$S.E(\bar{x}) = \frac{s}{\sqrt{n}}$$

Significance (Dice Example)

- Suppose that the outcome of the die throw experiment was a sum of 401. In previous lectures, a simulation study found that only in approximately 1.75% of cases would a fair die yield this result.
- However, in the case of a crooked die (i.e. one that favours high numbers) this result would not be unusual.
- A reasonable interpretation of this experiment is that the die is crooked, but importantly the experiment doesn't prove it one way or the other.
- We will discuss the costs of making a wrong decision later (Type I and Type II errors).