PO91Q: Fundamentals in Quantitative Research Methods

Your Questions, Weeks 1-5

Dr Florian Linke
Florian.Linke@warwick.ac.uk

1. Beta distribution, what is it? How is it different from normal?
   - The Beta distribution is a continuous probability distribution defined on the interval [0, 1].
   - I have only used this to create a strongly skewed distribution for simulation purposes.

2. What is the difference between a t-test and χ^2-test?
   - The t-test is used for one-sample or two-sample comparisons of means. It assesses whether two values are statistically different from each other, assuming the data is approximately normally distributed.
   - The χ^2-test is used to examine the association between two categorical variables. It assesses whether the observed frequencies in each category differ significantly from the expected frequencies under the null hypothesis of independence.

3. If the df value is something that we don't have in the table, what value should we take? Like for numbers like 99, 96, etc.
   - Use the closest available df for an approximation. For example, for df = 99, use df = 100; for df = 96, use df = 95.
   - In the exam, the tables will have all necessary df values.

4. After we reject the null hypothesis or fail to reject, what happens next? Never understood the conclusion of it.
   - A rejecting of the null hypothesis suggests that there is sufficient evidence to support the alternative hypothesis, indicating a significant effect or difference. This effect is usually the effect in the theory we are testing.
   - Failing to reject the null hypothesis suggests that there is insufficient evidence to support the alternative hypothesis, indicating no significant effect or difference. This does not prove the null hypothesis is true, just that we do not have enough evidence against it.

5. $P \leq 0.05$ (How is this value determined?)
   - The threshold of $P \leq 0.05$ is a conventional standard in many scientific fields for determining statistical significance. It indicates that there is a 5% chance of observing the data, or something more extreme, if the null hypothesis is true.
   - To determine it we determine the right-tail probability beond the observed t-value (or other test statistic) under the null hypothesis distribution.

6. Type 1 and type 2 Errors: Where do we use them? And what's their use? How do we interpret them?
   - Type 1 Error (False Positive): This occurs when we incorrectly reject a true null hypothesis. It is denoted by alpha (α), which is the significance level we set (e.g., 0.05). We use this concept to control the rate of false positives in hypothesis testing.
   - Type 2 Error (False Negative): This occurs when we fail to reject a false null hypothesis. It is denoted by beta (β). We use this concept to understand the power of a test, which is the probability of correctly rejecting a false null hypothesis (1 - β).
   - Example: Suppose a researcher tests whether a new policy reduces poverty.
     - The null hypothesis is that the policy has no effect.
     - A Type 1 Error would mean concluding the policy reduced poverty when in reality it had no impact (false positive).
     - A Type 2 Error would mean concluding the policy had no effect when in reality it really did reduce poverty (false negative).
     - Policymakers typically want to avoid Type 1 Errors if implementing an ineffective policy is costly, but want to avoid Type 2 Errors if missing a beneficial reform is costly.

7. How to interpret with CI if it is statistically significant or not?
   - huh?

8. For a large sample size (eg n = 100), the t-distribution is extremely close to the standard normal distribution. So, should we assume that 95% confidence will always give a value of 1.96? I tried using the z distribution and got a value against 99 df - which was 1.984, but that changed my CI
   - So, technically the t-distribution is only equal to the normal distribution when $df = \infty$.
   - But you are correct in that the t-distribution approaches the normal distribution as the sample size increases.
   - So for df = 99, the t-value is 1.984, which is very close to the z-value of 1.96, but still different.

9. Still can't tell when we do t, when we do z and when we do χ^2.
    - For normal vs. t-distribution, see Figure Figure 1.
    - Use χ^2-tests for examining relationships between categorical variables or for goodness-of-fit tests.

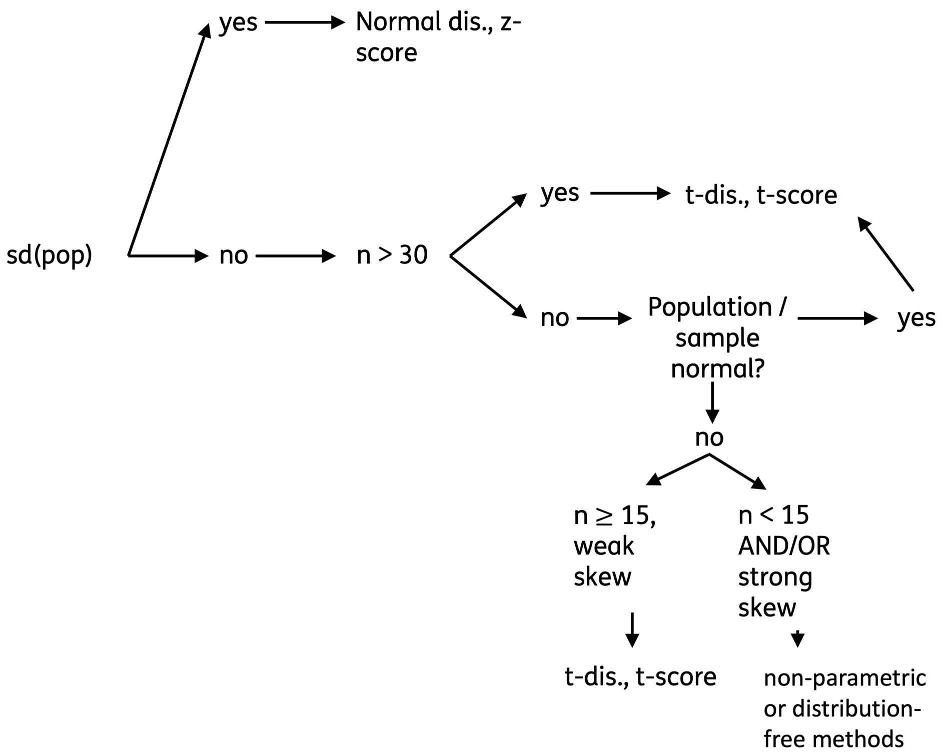

Figure 1: Distribution Choice, Decision Tree