Frequently Asked Questions

1. **What do you mean by a Randomized test?**
   Answer:
   A randomized test T is the one in which no test statistic is used. The decision about the rejection of the null hypothesis H_0 is taken, if it satisfies some predefined criterion. For example if it is decided that H_0 will be rejected if on tossing of a coin it falls with the head on the upper side and will be accepted if it falls with tail on the upper side . But randomized tests are rarely used.

2. **How do you define Non-Randomized test?**
   Answer:
   A test T of a hypothesis H is said to be nonrandomized if the hypothesis H0 is rejected on the basis that a test statistic belongs to the critical region C that is $\Phi(x1, x2 ...xn) \varepsilon C$

3. **How do you give the test function in randomized and non-randomized tests?**
   Answer:
   The sample space of observations Omega is partitioned into 2 regions, C and A. For a randomized test, we give the test procedure in terms of a function Φ as follows.
   $$\Phi(x) = P(\text{reject } H_0 \text{ when } X=x)$$
   For a non-randomized test with rejection region C, Φ for a region C is just its indicator function. That is,
   $$\psi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in C \\ 0 & \text{if } \mathbf{x} \notin C. \end{cases}$$
   A nonrandomized test procedure is a rule δ(X) that assigns two decisions to two disjoint subsets, C and A, of the range of T(X).
   We equate those two decisions with the real numbers 0 and 1, so ψ(X) is a real-valued function, where C is the critical region.
   If ψ(X) takes the value 0, the decision is not to reject; if ψ(X) takes the value 1, the decision is to reject. If the range of ψ(X) is {0, 1}, the test is a nonrandomized test.

4. **How do you give a test function when the outcome x is on the boundary of the critical region?**
   Answer:
   We will extend this, to allow for some different action (other than ‘reject’ and ‘accept’) if the outcome **x** is on the boundary of the critical region. The other action effectively is performing an auxiliary experiment such as tossing a coin with p (heads) $=$p; if heads results, reject H_0; if tails results, H_0 is accepted. The value of p is chosen to make the P (rejecting H_0) the desired value. More formally, for a test with critical region C and a value of $X=x_0$ on the boundary, we may define

$$\psi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in C \\ p & \text{if } \mathbf{x} = \mathbf{x_0} \\ 0 & \text{if } \mathbf{x} \neq \mathbf{x}_0 \text{ and } \mathbf{x} \notin C \end{cases}$$

where p ($0 < p < 1$) is appropriately chosen.

5. **Suppose X has a Poisson distribution with mean λ. A sample of size n = 10 is used to test H_0: λ=0.1 against H_1: λ> 0.1**
   Y= $\sum$Xi has a Poisson distribution with mean 1, the test is to reject H_0 for large values of Y. Suppose we wish to have a significance level of α = 0.05
   What should be the value of p?
   Given: P (Y≥3) =0.080 and P (Y≥4) =0.019 and the test function is,

$$\psi(y) = \begin{cases} 1 & \text{if } y \geq 4 \\ p & \text{if } y = 3 \\ 0 & \text{if } y < 3. \end{cases}$$

   Answer:
   Now p is found as follows.
   Size of the test =α

$$\begin{aligned} P(H_0 \text{ is rejected}) &= 1 \times P(Y \geq 4) + p \times P(Y = 3) + 0 \\ &= .019 \ + \ (.080 - .019)p \\ &= .019 + .061p \\ &= .05 \text{ if } p = 31/61. \end{aligned}$$

6. **Explain parametric tests with respect to randomized tests?**
   Answer:
   Both Parametric tests and randomized tests primary goal is to test some null hypothesis, although that null is distinctly different from what it would be with a parametric test.
   In parametric tests we randomly sample from one or more populations. We make certain assumptions about those populations, most commonly that they are normally distributed with equal variances. We establish a null hypothesis that is framed in terms of parameters. We use our sample statistics as estimates of the corresponding population parameters, and calculate a test statistic (such as a *t* test). We then refer that test statistic to the tabled sampling distribution of the statistic, and reject the null if our test statistic is extreme relative to the tabled distribution.

7. **What are the differences between parametric and randomized tests?**
   Answer:
   Randomization tests differ from parametric tests in almost every respect.

- We rarely think in terms of the populations from which the data came, and there is no need to assume anything about normality or homoscedasticity
- Our null hypothesis has nothing to do with parameters, but is phrased rather vaguely, as, for example, the hypothesis that the treatment has no effect on the how participants perform
  - That might be phrased a bit more precisely by saying that, under the null hypothesis, the score that is associated with a participant is independent of the treatment that person received
- Because we are not concerned with populations, we are not concerned with estimating (or even testing) characteristics of those populations.
- We do calculate some sort of test statistic; however we do not compare that statistic to tabled distributions
  - Instead, we compare it to the results we obtain when we repeatedly randomize the data across the groups, and calculate the corresponding statistic for each randomization

- Even more than parametric tests, randomization tests emphasize the importance of random assignment of participants to treatments

8. **When do we use randomized tests?**
   Answer:
   The randomization test of independence is used when you have two nominal variables. A data set like this is often called an "R×C table," where R is the number of rows and C is the number of columns. The randomization test is more accurate than the chi-squared tests or G-test of independence when the expected numbers are small.
   Fisher's exact test would be just as good as a randomization test, but there may be situations where the computer program you're using can't handle the calculations required for the Fisher's test.

9. **What are the advantages of Randomized tests?**
   Answer:
   A randomization test is not a different statistical test but a different, and always valid, method of determining statistical significance. The familiar *t*-test and *F*-test can be carried out by data permutation without any parametric assumptions being fulfilled. A particular advantage of this method is that unbalanced designs and missing values are easily accommodated. Even with only a small number of subjects the number of permutations will be large and a computer is necessary if the randomization test is to be of practical value. To make this method of determining statistical significance generally available an interactive microcomputer program, forming a comprehensive package for the design and analysis of experiments, has been prepared.

10. **What do you mean by a p- value?**
    Answer:
    The *p*-value is the probability of the test statistic being at least as extreme as the one observed given that the null hypothesis is true. A small *p*-value is an indication that the null hypothesis is false. It is the least value of α for which a null hypothesis is rejected.

11. **How do you take decision about the hypothesis based on p-value?**
    Answer:
    Although there is often confusion, the p-value is not the probability of the null hypothesis being true, nor is the p-value the same as the Type I error rate. Traditionally, one rejects the null hypothesis if the *p*-value is less than or equal to the significance level or Type I error; the connection is that a hypothesis test that rejects the null hypothesis for all samples that have a p-value less than α will have a Type I error of α. A significance level of 0.05 would deem extraordinary any result that is within the most extreme 5% of all possible results under the null hypothesis. In this case a p-value less than 0.05 would result in the rejection of the null hypothesis at the 5% (significance) level.

12. **What are the factors which may lead to misunderstandings about p-value**
    Answer:
    The data obtained by comparing the p-value to a significance level will yield one of two results: either the null hypothesis is rejected, or the null hypothesis cannot be rejected at that significance level (which however does not imply that the null hypothesis is true). A small p-value that indicates statistical significance does not indicate that an alternative hypothesis is correct.
    There are several common misunderstandings about p-values.

- The p-value is not the probability that the null hypothesis is true. In fact, frequentist statistics not, and cannot, attach probabilities to hypotheses
- The p-value is not the probability that a finding is "merely a fluke." As the calculation of a p-value is based on the assumption that a finding is the product of chance alone, it patently cannot also be used to gauge the probability of that assumption

being true. This is different from the real meaning which is that the p-value is the chance of obtaining such results if the null hypothesis is true

- The p-value is not the probability of falsely rejecting the null hypothesis
- The significance level of the test (denoted as α) is not determined by the p-value
- The significance level of a test is a value that should be decided upon by the agent interpreting the data before the data are viewed, and is compared against the p-value or any other statistic calculated after the test has been performed. (However, reporting a p-value is more useful than simply saying that the results were or were not significant at a given level, and allows readers to decide for themselves whether to consider the results significant.)
- The p-value does not indicate the size or importance of the observed effect (compare with effect size). The two do vary together however – the larger the effect, the smaller sample size will be required to get a significant p-value.

13. Briefly explain the conventions applied in p-value?

Answer:

Small p values provide evidence against the null hypothesis because they say the observed data are unlikely when the null hypothesis is true. We apply the following conventions:

- When p value > .10 → the observed difference is "not significant"
- When p value ≤ .10 → the observed difference is "marginally significant"
- When p value ≤ .05 → the observed difference is "significant"
- When p value ≤ .01 → the observed difference is "highly significant"

14. Explain the pocedure to interpret P value?

Answer:

It was common in the past for researchers to classify results as statistically 'significant' or 'non-significant', based on whether the *P* value was smaller than some prespecified cut point, commonly 0.05. This practice is now becoming increasingly obsolete, and the use of exact *P* values is much preferred. This is partly for practical reasons, because the increasing use of statistical software renders calculation of exact *P* values increasingly simple as compared with the past when tabulated values were used.

P- Value answers the question "If the null hypothesis were true, what is the probability of observing the current data or data that is more extreme?"

The use of "significant" in this context means, "the observed difference is not likely due to chance."

P-value answers the question: What is the probability of the observed test statistic when H_0 is true?

Thus, smaller and smaller P-values provide stronger and stronger evidence against H_0

15. How do you convert the Z statistic to P value?

Answer:

For H1: $\mu > \mu_0 \Rightarrow P = \Pr(Z > z_{stat})$ = right-tail beyond z_{stat}

For H1: $\mu < \mu_0 \Rightarrow P = \Pr(Z < z_{stat})$ = left tail beyond z_{stat}

For H1: $\mu \neq \mu_0 \Rightarrow P = 2 \times$ one-tailed P-value