Typical Steps for Statistical Hypothesis Testing

December 5, 2018

1. Observations
   $X_1, \ldots, X_n$ *i.i.d. from* $N(\mu, 1)$.

2. Specify the null and alternative hypotheses
   $H_0 : \mu = \mu_0$ *v.s.* $H_1 : \mu \neq \mu_0$ *(many other choices exist)*

3. Design a test statistic T and derive its sampling distribution under H_0
   In this example, the relevant test statistic is
   $$T = \frac{\bar{X} - \mu_0}{1/\sqrt{n}}.$$
   Due to the fact that $\bar{X} \sim N(\mu_0, n^{-1})$ *under* H_0*, then the test statistic* $T \sim N(0, 1)$.

4. Calculate the value of test statistic with observed data T_{obs}, compare with the critical value
   Note that the test statistics T *is itself a random variable, while its observed value* T_{obs} *is a scalar once observations are given.*
   In this example, let the significance level $\alpha = 0.05$*, then the critical value is* 1.96*, which means* $P(|T| > 1.96) = 0.05$*. If the observed test statistic* T_{obs} *is outside of interval* $[-1.96, 1.96]$*, then reject* H_0*; otherwise, fail to reject the null.*

5. Equivalent to the above step, we usually calculate the p-value and compare it with the significant level α
   The meaning of p-value is the probability, under the null hypothesis H_0*, of obtaining a test statistic which is more or equally extreme than the one we have observed.*
   In this example, the p-value is $P_{H_0}(|T| > |T_{obs}|)$*. When compared with* $\alpha = 0.05$*, it is equivalent to compare* $|T_{obs}|$ *and the critical value* 1.96*.*

6. Decision: reject H_0 or fail to reject H_0.
   Note that fail to reject H_0 does not mean we accept H_0. Along the steps of hypothesis testing, we are assuming H_0 is true. The terminology of *fail to reject* means there is a lack of evidence against H_0. The terminology of *accept* H_0 is apparently a stronger conclusion, and it is improper used in hypothesis testing.