Motivation for Hypotheses Testing

- Business: Will an investment in a mutual fund yield annual returns greater than desired value? (based on past performance of the fund)
- Medical: Is the incidence of diabetes greater among males than females?
- Social: Are women more likely to change mobile service provider than men?
- Engineering: Has the efficiency of the pump (η) decreased from its original value due to aging?

Hypotheses Testing

- The hypotheses is generally converted to a test of the mean or variance parameter of a population (or differences in means or variances of populations)
- A hypothesis a statement or postulate about the parameters of a distribution (or model)
  - Null hypothesis H_0 : The default or *status quo* postulate that we wish to reject if the sample set provides sufficient evidence (eg. $\eta = \eta_0$)
  - Alternative hypothesis H_1 : The alternative postulate that is accepted if the null hypothesis is rejected (eg. $\eta < \eta_0$)

Hypotheses Testing Procedure

- Identify the parameter of interest (mean, variance, proportion) which you wish to test
- Construct the null and alternative hypotheses
- Compute a test statistic which is a function of the sample set of observations
- Derive the distribution of the test statistic under the null hypothesis assumption
- Choose a test criterion (threshold) against which the test statistic is compared to reject/not reject the null hypothesis

Hypotheses Testing Procedure

- No hypotheses test is perfect. There are inherent errors since it is based on observations which are random
- The performance of a hypotheses test depends on
  - Extent of variability in data
  - Number of observations (Sample size)
  - Test statistic (function of observations)
  - Test criterion (threshold)

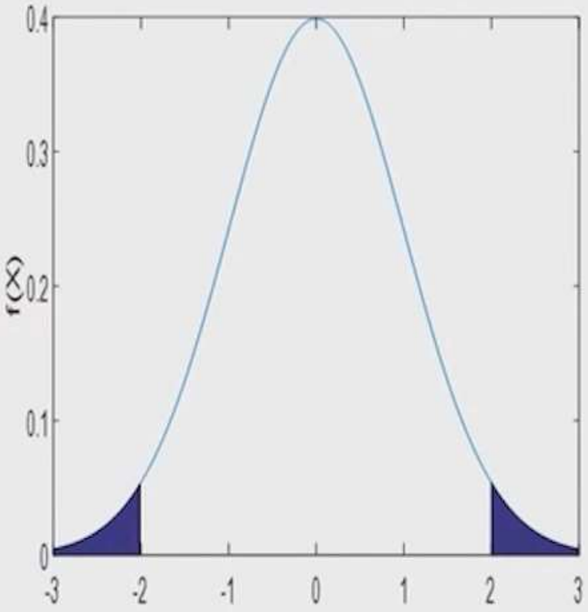
Two-sided and one-sided tests

- Two sided test

  $H_0: \mu = 0$

  $H_1: \mu \neq 0$

- Test statistic standard normal RV z



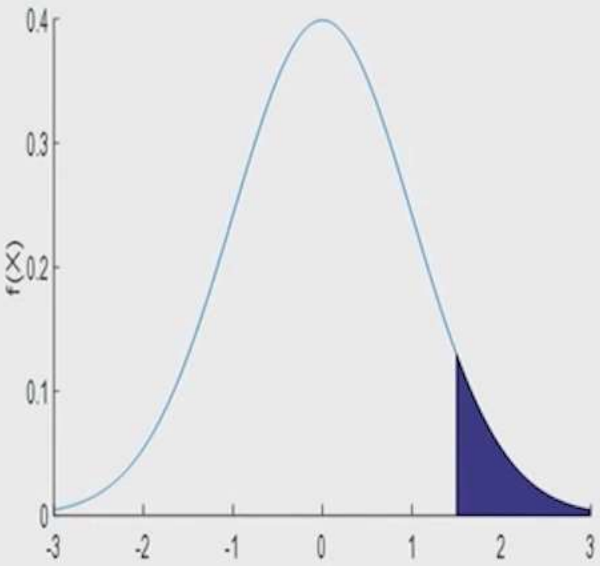
- Reject H_0 if $z \leq -2$ or $z \geq 2$

- One sided test

  $H_0: \mu = 0$

  $H_1: \mu > 0$

- Test statistic standard normal RV z



- Reject H_0 if $z \geq 1.5$

Errors in Hypotheses Testing

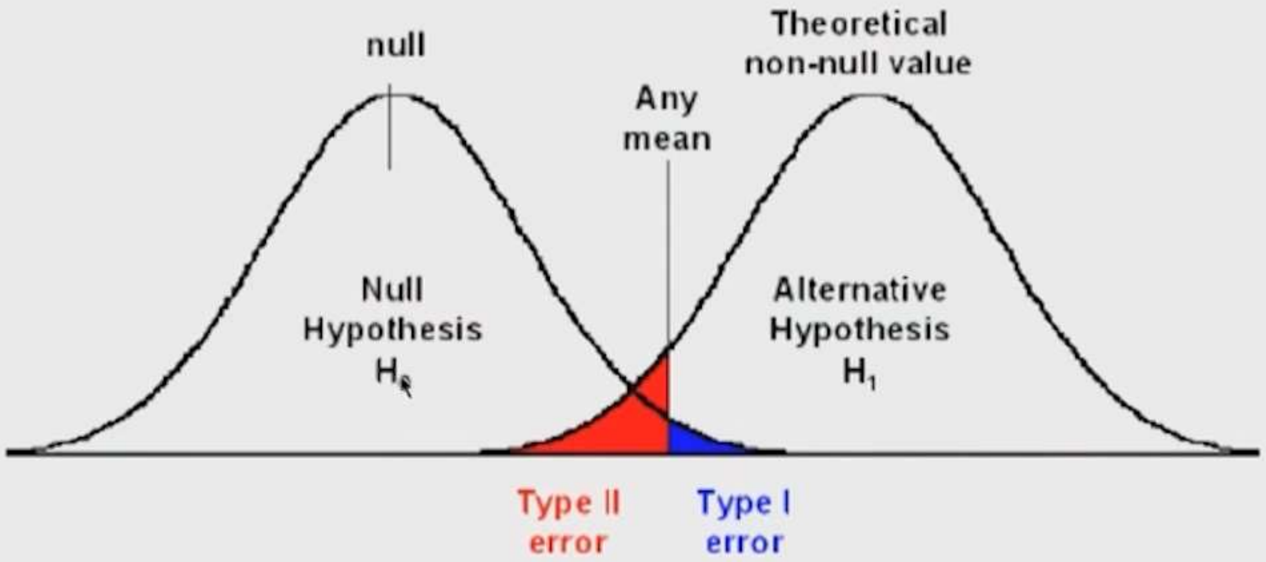
- Two Types of errors (Type I and Type II)

Decision → Truth ↓	H_0 is not rejected	H_0 is rejected
H_0 is true	Correct Decision $Pr = 1 - \alpha$	Type I error $Pr = \alpha$
H_1 is true	Type II error $Pr = \beta$	Correct Decision $Pr = 1 - \beta$

  - Typically the Type 1 error probability α (also called as level of significance of the test) is controlled by choosing the criterion from the distribution of the test statistic under the null hypothesis

Errors in Hypotheses Testing

- Type I and Type II error probabilities



- Statistical test Power = 1 – Type II error probability
- Trade-off : If we decrease Type I error probability, then Type II error probability will increase

Test for Mean : Solid Propellant example

For a given application the burning rate of a solid propellant should be 50 cm/s.

- 25 samples of the solid propellant are taken and their burning rate noted. The average burning rate is computed to be 51.3 cm/s. The standard deviation in the burning rate is known to be 2 cm/s
- Null hypothesis : $\mu = 50$ cm/s
- Alternative hypothesis : $\mu \neq 50$ cm/s (lower or higher burning rate propellants are both unsatisfactory) – Two sided test
- Test statistic $z = \frac{\bar{x}-50}{2/\sqrt{25}} \sim \mathcal{N}(0,1)$; $z = 3.25$
- Critical value for $\alpha = 0.05$ is ± 1.96
- Decision: Reject null hypothesis

Test for Differences in Means : Training example

Two groups of teachers of similar capabilities are trained by two methods A and B. Is Method B more effective than Method A?

- 10 teachers in each group. Average scores and standard deviation of scores after training are Group 1: $\bar{x}_1 = 70, s_1 = 3.3665$ Group 2: $\bar{x}_1 = 74, s_1 = 5.3955$
- Null hypothesis : $\mu_1 - \mu_2 = 0$
- Alternative hypothesis : $\mu_1 - \mu_2 < 0$ – one sided test
- Test statistic (assuming unknown but equal variances for two groups)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{N_1} + \frac{s_p^2}{N_2}}} \sim t_{N_1+N_2-2}; \quad S_p = \frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1+N_2-2} \quad t = -1.989$$

- Critical value for $\alpha = 0.05$ is -1.73
- Decision: Reject null hypothesis (Method B is better)

Test for Differences in Variances : Process Yields

The variability in yields from two different processes are to be compared to decide whether they are identical or not

- 50 samples for each process taken. Yield variances are found to be $s_1^2 = 2.05$ and $s_2^2 = 7.64$
- Null hypothesis : $\frac{\sigma_1^2}{\sigma_2^2} = 1$
- Alternative hypothesis : $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$ – two sided test
- Test statistic (assuming unknown but equal variances for two groups)

  $$\mathrm{f} = \frac{s_1^2}{s_2^2} \sim \mathrm{F}(N_1 - 1, N_2 - 1);\ \mathrm{f} = 0.27$$

- Critical value for α = 0.025 is 0.567 and α = 0.975 is 1.762
- Decision: Reject null hypothesis (Process 2 has higher variability)

Summary of useful hypotheses tests

Type of test	Characteristic	Example	Application
z-test	Sum of independent normal variables	Test for a mean or comparison between two group means (variance known)	Test coefficients of a regression model
t-test	Ratio of a standard normal variable and chi-square variables with p degrees of freedom	Test for a mean or comparison between two group means (variance unknown)	Test coefficients of a regression model
chi-square test (p degrees of freedom)	Sum of p independent standard normal variables	Test for variance	Test quality of regression model
F-test (p_1 and p_2 degrees of freedom)	Ratio of two chi-square variables	Test for comparing variances of two groups	Choose between regression models having different number of parameters