# AE 102: Data Analysis and Interpretation

January - April 2016
Department of Aerospace Engineering

## Sample

- $X_1, X_2, ... X_n$ are independent
- with common distribution $F$
- Parametric inference: assume distribution $F$ but without specifics
- Non-parametric inference

## Sample mean

- Consider a population with given mean and variance
- Take a sample of $n$ elements

$$\bar{X} = \frac{X_1 + ... + X_n}{n}$$

- What is $E[\bar{X}]$?
- What is $\mathrm{Var}(\bar{X})$?

## The Central Limit Theorem

- If $X_1, X_2, ... X_n$ are iid each with $\mu$ and $\sigma$
- then for $n$ large,
- $\sum_i X_i$ is approximately normal with mean $n\mu$ and variance $n\sigma^2$
- So how large should $n$ be?

## Sample Variance

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

- Compute $E[S^2]$

## Sample Variance

Using:

- $\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- $E[X^2] = Var(X) + E[X]^2$
- Show that: $E[S^2] = \sigma^2$

## Sampling from a Normal Population

- Population is normal with $\mu, \sigma$
- How are $\bar{X}, S^2$ distributed?

## Distribution of Sample Mean

- $\bar{X}$ is normally distributed
- $E[\bar{X}] = \mu$
- $Var[\bar{X}] = \sigma^2/n$

## Joint Distribution of mean and variance

- $\bar{X}$ and $S^2$ are independent
- $(n-1)S^2/\sigma^2$ is $\chi^2_{n-1}$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \mu)^2 - n(\bar{x} - \mu)^2$$

## Distribution of $S^2$

$$\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} - \left[\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}\right]^2$$

---

- LHS: $\chi^2_n$
- RHS, second term, $\chi^2_1$

## Corollary

- Sample of $n$ items from a normal population

$$\sqrt{n}\frac{(\bar{X} - \mu)}{S} \sim t_{n-1}$$

---

- Recall: $\sqrt{n}Z/\sqrt{\chi^2_n}$