# Independece

$X$ and $Y$ are **independent** iff

- $f(x,y) = f(x)f(y)$
- $F(x,y) = F(x)F(y)$
- $E(XY) = E(X)E(Y)$
- $m_{X,Y}(t_1,t_2) = m_X(t_1)m_Y(t_2)$
- If $X$ and $Y$ are independent then so are $U = g(X)$ and $V = h(Y)$. That is, functions of independent random variables are also independent.

1. **Normal Distribution** If $X_i's$ are i.i.d, $N(\mu, \sigma^2)$ random variable, then
   - $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
   - $Z = \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) \sim N(0,1) \Rightarrow Z^2 = \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi_1$
   - $\bar{X}$ and $S^2$ are independent
   - $\frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{(n-1)}$
   - If $Z_i's$ are i.i.d, $N(0,1)$ random variable, then
     - $Z_i^2 \sim \chi_1^2$
     - $\sum_{i=1}^n Z_i^2 \sim \chi_n^2$

2. **t-Distribution** Given $Z \sim N(0,1)$, $W \sim \chi_v^2$, and $Z \perp\!\!\!\perp W$, then
   - $T = \frac{Z}{\sqrt{W/v}} \sim t_v$
   - $T^2 \sim F(1, n)$

3. **F Distribution** If $W_1 \sim \chi_{v_1}^2$, $W_2 \sim \chi_{v_2}^2$, $W_1 \perp\!\!\!\perp W_2$, then
   - $F = \frac{W_1/(v_1)}{W_2/(v_2)} \sim F(v_1, v_2)$

4. **Gamma Distribution** If $X \sim Gam(\alpha_1, \beta)$, $Y \sim Gam(\alpha_2, \beta)$, $X \perp\!\!\!\perp Y$, then
   - If $c \in R^+$, then $cX \sim Gam(\alpha_1, c\beta)$
   - $X + Y \sim Gam(\alpha_1 + \alpha_2, \beta)$
   - $\text{Exp}(\beta) = Gam(1, \beta)$
   - $\chi_n^2 = Gam(\frac{n}{2}, 2)$

# Unbiased Estimator

- Let $\hat{\theta}$ be an estimator for a parameter $\theta$. Then $\hat{\theta}$ is an *unbiased estimator* if
$$\mathbb{E}[\hat{\theta}] = \theta.$$
  - If $\mathbb{E}[\hat{\theta}] \neq \theta.$, then $\hat{\theta}$ is an biased estimator.
- Bias
$$\mathbb{B}(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta.$$
  - If $\hat{\theta}$ is unbiased, then $\mathbb{B}(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta = 0$.
- How do we choose the best estimator among unbiased estimators? $\Rightarrow$ *MSE*

# Mean Square Error (MSE)

- MSE of a point estimator $\hat{\theta}$ is

$$\mathrm{MSE}_\theta(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

- The MSE can be expressed in terms of the variance and the bias of a estimator as:

$$\mathrm{MSE}_\theta(\hat{\theta}) = \mathbb{V}[\hat{\theta}] + [\mathbb{B}(\hat{\theta})]^2.$$

That is, for an unbiased estimator $\mathbb{B}(\hat{\theta}) = 0$

$$\mathrm{MSE}_\theta(\hat{\theta}) = \mathbb{V}[\hat{\theta}].$$

# Example 1

- Let $X_1, \ldots, X_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ random variables, where $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$ are unknown.
- $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$.
- $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$.

a) $\bar{X}$, $X_1$ and $\frac{X_1+X_2}{2}$ are unbiased estimators of $\mu$.

b) $S^2$ and $\frac{(X_1-X_2)^2}{2}$ are unbiased estimators of $\sigma^2$.
c) $S_n^2$ is a biased estimator of $\sigma^2$.

1. Find $\mathbb{B}(S_n^2)$.
2. Find $\mathrm{MSE}_{\sigma^2}(S_n^2)$.
3. Find $\mathbb{B}(S^2)$.
4. Find $\mathrm{MSE}_{\sigma^2}(S^2)$
5. Which estimator is better?