

**Intro: Random Experiment:** An outcome for each run.

**Sample Space  $\Omega$ :** Set of all possible outcomes.

**Event:** Subsets of  $\Omega$ .

**Prob. of Event  $A$ :**  $P(A) = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } \Omega}$

**Axioms:**  $P(A) \geq 0 \forall A \in \Omega$ ,  $P(\Omega) = 1$ ,

If  $A \cap B = \emptyset$ , then  $P(A \cup B) = P(A) + P(B) \forall A, B \in \Omega$

**Cond. Prob.**  $P(A|B) = \frac{P(A \cap B)}{P(B)}$

\*  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

**Independence:**  $P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A)P(B)$

**Total Prob. Thm:** If  $H_1, H_2, \dots, H_n$  form a partition of  $\Omega$ , then  $P(A) = \sum_{i=1}^n P(A|H_i)P(H_i)$ .

**Bayes' Rule:**  $P(H_k|A) = \frac{P(H_k \cap A)}{P(A)} = \frac{P(A|H_k)P(H_k)}{\sum_{i=1}^n P(A|H_i)P(H_i)}$

\*Posteriori:  $P(H_k|A)$ , Likelihood:  $P(A|H_k)$ , Prior:  $P(H_k)$

**1 RV: CDF:**  $F_X(x) = P[X \leq x]$

**PMF:**  $P_X(x_j) = P[X = x_j] \ j = 1, 2, \dots$

**PDF:**  $f_X(x) = \frac{d}{dx} F_X(x)$

\*  $P[a \leq X \leq b] = \int_a^b f_X(x) dx$  IS THIS CORRECT?

**Cond. PMF:**  $P_X(x|A) = P[X = x|A] = \frac{P[X=x, A]}{P[A]}$  IS THIS CORRECT?

**Cond. PDF:**  $f_X(x|A) = \frac{f_{X,A}(x,a)}{f_A(a)}$  IS THIS CORRECT?

**Exp.:**  $E[h(X)] = \int_{-\infty}^{\infty} h(x)f_X(x)dx \mid \sum_{k=-\infty}^{\infty} h(k)f_X(k)$

**Variance:**  $\text{Var}[X] = E[(X - \mu)^2] = E[X^2] - E[X]^2$

**Cond. Exp.:**  $E[X|A] = \int_{-\infty}^{\infty} x f_X(x|A) dx$

**2 RVs: Joint PMF:**  $P_{X,Y}(x, y) = P[X = x, Y = y]$

**Joint PDF:**  $f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$

\*  $P[(X, Y) \in A] = \int \int_{(x,y) \in A} f_{X,Y}(x, y) dx dy$

**Expectation:**  $E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$

**Correlation:**  $E[XY]$

**Covariance:**  $\text{Cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - E[X]E[Y]$

**Correlation Coeff.:**  $\rho_{X,Y} = E \left[ \left( \frac{X - \mu_X}{\sigma_X} \right) \left( \frac{Y - \mu_Y}{\sigma_Y} \right) \right] = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y}$

**Marginal PMF:**  $P_X(x) = \sum_{j=1}^{\infty} P_{X,Y}(x, y_j)$

**Marginal PDF:**  $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$

**Conditional PMF:**  $P_{X|Y}(x|Y) = P[X = x|Y = y] = \frac{P_{X,Y}(x, y)}{P_Y(y)}$

**Conditional PDF:**  $f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$

**Bayes' Rule**  $f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{f_{X|Y}(x|y)f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x|y')f_Y(y') dy'}$

\*  $P_{Y|X}(y|x) = \frac{P_{X,Y}(x, y)}{P_X(x)} = \frac{P_{X|Y}(x|y)P_Y(y)}{\sum_{j=1}^{\infty} P_{X|Y}(x|y_j)P_Y(y_j)}$

**Independent:**  $f_{X|Y}(x|y) = f_X(x) \forall y \Leftrightarrow f_{X,Y}(x, y) = f_X(x)f_Y(y)$

\*If independent, then uncorrelated.

**Uncorrelated:**  $\text{Cov}[X, Y] = 0 \Leftrightarrow \rho_{X,Y} = 0$

**Orthogonal:**  $E[XY] = 0$

**Conditional Expectation:**  $E[Y] = E[E[Y|X]]$  or  $E[E[h(Y)|X]]$

\*  $E[E[Y|X]]$  w.r.t.  $X \mid E[Y|X]$  w.r.t.  $Y$ .

**Estimation:** Estimate unknown parameter  $\theta$  from  $n$  i.i.d. measurements  $X_1, X_2, \dots, X_n$ ,  $\hat{\theta}(\underline{X}) = g(X_1, X_2, \dots, X_n)$

**Estimation Error:**  $\hat{\theta}(\underline{X}) - \theta$ .

**Unbiased:**  $\hat{\theta}(\underline{X})$  is unbiased if  $E[\hat{\theta}(\underline{X})] = \theta$ .

\* **Asymptotically unbiased:**  $\lim_{n \rightarrow \infty} E[\hat{\theta}(\underline{X})] = \theta$ .

**Consistent:**  $\hat{\theta}(\underline{X})$  is consistent if  $\hat{\theta}(\underline{X}) \rightarrow \theta$  as  $n \rightarrow \infty$  or  $\forall \epsilon > 0, \lim_{n \rightarrow \infty} P[|\hat{\theta}(\underline{X}) - \theta| < \epsilon] \rightarrow 1$ .

**Sample Mean:**  $M_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

\* Given a sequence of i.i.d. RVs,  $X_1, X_2, \dots, X_n$ ,  $M_n$  is unbiased and consistent.

**Weak Law of Large #s**  $\lim_{n \rightarrow \infty} P[|M_n - \mu| < \epsilon] = 1 \ \forall \epsilon > 0$ .

**Maximum Likelihood Estimation:** Choose parameter  $\theta$  that is most likely to generate the obs.  $x_1, x_2, \dots, x_n$ .

\* Discrete  $X$ :  $\hat{\theta} = \arg \max_{\theta} P_{\underline{X}}(\underline{x}|\theta) \stackrel{\log}{=} \arg \max_{\theta} \sum_{i=1}^n \log P_X(x_i|\theta)$

\* Cont.  $X$ :  $\hat{\theta} = \arg \max_{\theta} f_{\underline{X}}(\underline{x}|\theta) \stackrel{\log}{=} \arg \max_{\theta} \sum_{i=1}^n \log f_X(x_i|\theta)$