BIVARIATE RANDOM VARIABLES AND 
JOINT DISTRIBUTION FUNCTIONS

CHAPTER 3.1-3.3
A pair of RV \((X,Y)\) that associates two real numbers with every element in \(S\)

- Two-dimensional random vector

Function that maps outcome \(\xi\) to a point in the \((x,y)\)-plane

- Range of \((X,Y)\)
  \[R_{XY} = \{(x,y); \xi \in S \text{ and } X(\xi) = x, Y(\xi) = y\}\]
BIVARIATE RV TYPES

- Bivariate discrete RV – both $X, Y$ discrete
- Bivariate continuous RV – both $X, Y$ continuous
- Bivariate mixed RV – one discrete other continuous

- In this class will primarily focus on either bivariate discrete or continuous, not mixed
Joint Distribution Functions (CDF)

- $F_{XY}(x, y) = P(X \leq x, Y \leq y) = P(A \cap B)$
  - Event A: $(X \leq x)$; Event B: $(Y \leq y)$

- Formally, event $(X \leq x, Y \leq y) = \text{event } (A \cap B)$
  - $A = \{\xi \in S; X(\xi) \leq x\}$
    - $P(A) = F_X(x)$
  - $B = \{\xi \in S; Y(\xi) \leq y\}$
    - $P(B) = F_Y(y)$

- Independent RV
  - $F_{XY}(x, y) = F_X(x)F_Y(y) = P(A)P(B)$

- Properties – same general idea as for single RV
MARGINAL DISTRIBUTION

- Given joint CDF,
  - \( F_X(x) = F_{XY}(x, \infty) \)
  - \( F_Y(y) = F_{XY}(\infty, y) \)
- These are the distribution taking into account all values of the other RV
  - E.g. marginalizing/removing the effects/dependence on one variable
- Result comes from observation
  - \( \lim_{y \to \infty} (X \leq x, Y \leq y) = (X \leq x, Y \leq \infty) = (X \leq x) \)
  - The condition \( (Y \leq \infty) \) is always satisfied
JOINT PMF, JOINT PDF, AND CONDITIONAL DISTRIBUTIONS

CHAPTER 3.4-3.6
JOINT PMF

- Let \((X, Y)\) be discrete RV with values \((x_i, y_j)\) for an allowable set of integers \(i, j\)
  - \(p_{XY}(x_i, y_j) = P(X = x_i, Y = y_j)\)
- Properties
  1) \(0 \leq p_{XY}(x_i, y_j) \leq 1\)
  2) \(\sum_{x_i} \sum_{y_j} p_{XY}(x_i, y_j) = 1\)
  3) \(P[(X, Y) \in A] = \sum \sum_{(x_i, y_j) \in R_A} p_{XY}(x_i, y_j)\)
    - Points \((x_i, y_j)\) in \(R_A\) are in range space corresponding to event \(A\)
- CDF from PMF
  - \(F_{XY}(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p_{XY}(x_i, y_j)\)
MARGINAL PMF

- \( P(X = x_i) = p_X(x_i) = \sum_{y_j} p_{XY}(x_i, y_j) \)
  - Summation is over all possible \( Y = y_j \) values
  - Marginalize by removing influence of RV \( Y \)
- \( P(Y = y_j) = p_Y(y_j) = \sum_{x_i} p_{XY}(x_i, y_j) \)
- Independence:
  - \( p_{XY}(x_i, y_j) = p_X(x_i)p_Y(y_j) \)
**JOINT PDF**

- \((X,Y)\) is a continuous bivariate RV with CDF \(F_{XY}(x,y)\)
  - \(f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)\)
  - \(F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(\xi, \eta) d\eta d\xi\)

- Properties:
  1) \(f_{XY}(x,y) \geq 0\)
  2) \(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y) dxdy = 1\)
  4) \(P[(X,Y) \in A] = \int \int_{R_A} f_{XY}(x,y) dxdy\)
  5) \(P(a < X \leq b, c < Y \leq d) = \int_{c}^{d} \int_{a}^{b} f_{XY}(x,y) dxdy\)
MARGINAL PDF

- $F_X(x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(\xi, \eta) d\eta d\xi$
  - Integrate/marginalize over full range/all values of $y$
- $f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{\infty} f_{XY}(x, \eta) d\eta = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$
- $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx$
- Independence:
  - $F_{XY}(x, y) = F_X(x)F_Y(y)$
  - $f_{XY}(x, y) = f_X(x)f_Y(y)$
(X, Y) discrete bivariate RV with joint PMF \( p_{XY}(x_i, y_j) \)

\[ p_{Y|X}(y_j|x_i) = \frac{p_{XY}(x_i, y_j)}{p_X(x_i)}, \quad p_X(x_i) > 0 \]

Conditional PMF of Y given \( X = x_i \) → probability of \( Y = y_j \) knowing that \( X = x_i \)

Properties

1) \( 0 \leq p_{Y|X}(y_j|x_i) \leq 1 \)

2) \( \sum_j p_{Y|X}(y_j|x_i) = 1 \)

Independence

\[ p_{Y|X}(y_j|x_i) = p_Y(y_j) \quad \text{and} \quad p_{X|Y}(x_i|y_j) = p_X(x_i) \]
CONDITIONAL PDF

- \((X,Y)\) continuous bivariate RV with joint PMF \(f_{XY}(x,y)\)
  - \(f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}, \quad f_X(x) > 0\)
  - Conditional PDF of \(Y\) given \(X (= x)\)

- Properties
  1) \(f_{Y|X}(y|x) \geq 0\)
  2) \(\int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = 1\)

- Independence
  - \(f_{Y|X}(y|x) = f_Y(y) \quad \text{and} \quad f_{X|Y}(x,y) = f_X(x)\)
COVARIANCE/CORRELATION COEFFICIENT AND CONDITIONAL MEANS/VARIANCES

CHAPTER 3.7-3.8
\((k,n)^{th}\) MOMENT

- \(m_{kn} = E[X^kY^n]\)
  - Discrete: \(m_{kn} = \sum y_j \sum x_i x_i^k y_j^n p_{XY}(x_i, y_j)\)
  - Continuous: \(m_{kn} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^n f_{XY}(x, y)dx\,dy\)
- Note: \(m_{10} = E[X] = \mu_X\) and \(m_{01} = E[Y] = \mu_Y\)

\[
\mu_X = \sum_{y_j} \sum_{x_i} x_i y_j^0 p_{XY}(x_i, y_j) \\
= \sum_{x_i} x_i \sum_{y_j} p_{XY}(x_i, y_j) = \sum_{x_i} x_i p_X(x_i) \\
\text{marginalize}\]

Note: \(m_{10} = E[X] = \mu_X\) and \(m_{01} = E[Y] = \mu_Y\)
Measure of relationship between two RV

- $m_{11} = E[XY]$
- Measure away from independence (statistical)

If $E[XY] = 0$, then X and Y are orthogonal

- Note: orthogonal does not mean independent
- Think of an inner product in RV space $\rightarrow$ 90 degree angle vs. statistical independence

Note: “correlation does not imply causation”

- Just because two variables are correlated, does not mean that one causes the other
- E.g. increase in ice cream sales correlated with increase shark attacks. Probably not ice cream causing shark attacks but that ice cream and shark attacks happen more often during the summer
COVARIANCE

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

- If $\text{Cov}(X,Y) = 0 \Rightarrow X \text{ and } Y \text{ uncorrelated}$
  
  - $E[XY] = E[X]E[Y]$

- Note that independent RV are uncorrelated but uncorrelated does not imply independent
PEARSON’S CORRELATION COEFFICIENT

- Measure of **linear** dependence between $X, Y$

- $\rho(X, Y) = \rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

- $|\rho_{XY}| \leq 1$
Discrete

Mean (expectation)
- $\mu_{Y|x_i} = E[Y|x_i] = \sum_{y_j} y_j p_{Y|x}(y_j|x_i)$

Variance
- $\sigma^2_{Y|x_i} = Var(Y|x_i) = E[(Y - \mu_{Y|x_i})^2|x_i]$
  = $\sum_{y_j} (y_j - \mu_{Y|x_i})^2 p_{Y|x}(y_j|x_i)$
  = $E[Y^2|x_i] - E^2[Y|x_i]$ 

Note: these values are a function of $x_i$ and do not depend on $Y$
- Defined for different $x_i$ values

Continuous

Mean
- $\mu_{Y|x} = E[Y|x] = \int_{-\infty}^{\infty} y f_{Y|x}(y|x) dy$

Variance
- $\sigma^2_{Y|x_i} = Var(Y|x)$
  = $E[(Y - \mu_{Y|x})^2|x]$
  = $\int_{-\infty}^{\infty} (y - \mu_{Y|x})^2 f_{Y|x}(y|x) dy$
N-VARIATE RVS AND SPECIAL DISTRIBUTIONS

CHAPTER 3.8-3.9
N-VARIATE RV

- Natural extension of bivariate discussion
- Give n-tuple of RVs \((X_1, X_2, \ldots, X_n)\) – n-dim random vector
  - Each \(X_i\) \(i = 1, 2, \ldots, n\) associates a real number to sample point \(\xi \in S\)

- We won’t really work beyond bivariate in class
  - Ex: Joint CDF \(F_{X_1X_2\ldots X_n}(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)\)
Just like with single RV, there are important distributions that show up in nature a lot

- Multinomial distribution – extension of binomial
- N-variate Normal distribution
MULTINOMIAL DISTRIBUTION

- Multinomial trial (extension of binomial)
  - 1) Experiment with $k$ possible outcomes that are mutually exclusive ($A_1, A_2, \ldots, A_k$)
  - 2) $P(A_i) = p_i; \quad i = 1, \ldots, k; \quad \sum_{i=1}^{k} p_i = 1$

- Multinomial RV
  - $(X_1, X_2, \ldots, X_n)$ with $X_i$ be RV denoting number of trials with result $A_i$
    - Count of number of each outcome
  - $p_{x_1 x_2 \ldots x_k}(x_1, \ldots, x_k) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$
    - Probability of combination of different outcomes
MULTINOMIAL EXAMPLE

- \( k \) different color balls in a bag \( \rightarrow p_i \) is the probability of color \( i \) to be drawn

- Select a ball at random and record the color then replace in bag

- Count of the colors at the end of the \( n \) ball draws is a multinomial RV

- Distribution tells the probability of seeing e.g. 1 white, 2 red, 3 blue, and 4 green balls
NORMAl DISTRIBUTION

- **Bivariate**
  \[ f_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y(1 - \rho^2)^{1/2}} \exp \left[ -\frac{1}{2} q(x, y) \right] \]
  \[ q(x, y) = \frac{1}{1 - \rho^2} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right] \]

- **N-variate**
  - Vector valued function (see book for details)
  \[ f_X(x) = \frac{1}{(2\pi)^{n/2} |\det K|^{1/2}} \exp \left[ -\frac{1}{2} (x - \mu)^T K^{-1} (x - \mu) \right] \]
  - Covariance matrix
  \[ K = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix} \quad \sigma_{ij} = \text{Cov}(X_i, X_j) \]
  - Note: covariance controls shape or orientation in bivariate case