Chapter 5 Notes: Multivar. Prob. Distr.

Discrete Random Vectors

- random vector is a collection of random variables such as $(X_1, X_2, ...)$ defined on the same sample space
- Discrete random vector: (X, Y) where X and Y are both discrete random variables
- joint distribution: X takes values in $\{x_1, x_2, ...\}$ and Y takes values in $\{y_1, y_2, ...\}$
- joint probability mass function: $p_{ij} = P(X = x_i, Y = y_j)$; then $p_{ij} \geq 0$ and $\Sigma_i \Sigma_j p_{ij} = 1$
- joint cumulative distribution function (cdf): for any x, y $\in \mathbb{R}$:

$$F(x, y) = P(X \leq x, Y \leq y) = \Sigma_{x_i \leq x} \Sigma_{y_i \leq y} P(X = x_i, Y = y_j)$$

Continuous Random Vector

- definition: (X, Y) where X and Y are both continuous random variables
- joint probability density function: a nonnegative function $f : \mathbb{R}^2 \to \mathbb{R}_+$ such that for any subset $B \subseteq \mathbb{R}^2$:

$$P((X, Y) \in B) = \int \int_B f(x, y) dx dy$$

- $\int \int_{\mathbb{R}^2} f(x, y) dx dy = 1$
- joint cumulative distribution function (cdf): for any $x, y \in \mathbb{R}$:

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv$$

- relation between density and CDF:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Uniform Distribution

- Uniform Distribution on region B: Let $B \subseteq \mathbb{R}^2$ and define c = {area of B}, then the uniform distribution on B has joint probability density function:

$$f(x, y) = \begin{cases} 1/c & \text{if } (x, y) \in B \\ 0 & \text{otherwise} \end{cases}$$

- Let (X, Y) be uniform on B, then for any $A \subseteq B$:

$$P((X, Y) \in A) = \frac{\text{Area of A}}{\text{Area of B}}$$

Independence

- X and Y are said to be independent if for any subsets $A, B \subseteq \mathbb{R}$:

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$$

- Let F be the joint CDF of (X, Y) and F_X, F_Y be the CDFs of X and Y; then X and Y are independent if and only if for every x, y $F(x, y) = F_X(x) F_Y(y)$
- if X and Y are discrete random variables then X and Y are independent if and only if for every x_i and y_j, $P(X = x_i, Y = y_j) = P(X = x_i) P(Y = y_j)$
- if X and Y are continuous random variables, then X and Y are independent if and only if for every x, y, $f(x, y) = f_X(x) f_Y(y)$ where f is the joint density of (X, Y) and f_x is the density of X and f_y is the density of Y
- if X and Y are continuous random variables, X and Y are independent if and only if the joint density f is of form $f(x, y) = g(x) h(y)$
- if X and Y are independent, then g(X) and h(Y) are also independent for any function g, h: $\mathbb{R} \to \mathbb{R}$

Sum of Independent Normals

- Let X be $N(\mu_1, \sigma_1^2)$ and Y be $N(\mu_2, \sigma_2^2)$
- if X and Y are independent, then for any constants $a, b \in \mathbb{R}$ aX+bY is also normal:

$$aX + bY \approx N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

Expected Values

- Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ b e any function, and (X, Y) any random vector:
- If (X, Y) is discrete, then $E[h(X, Y)] = \Sigma_i \Sigma_j h(x_i, y_j) P(X = x_i, Y = y_j)$
- If (X, Y) is continuous with joint density f, then

$$E[h(X, Y)] = \int \int_{\mathbb{R}^2} h(x, y) f(x, y) dx dy$$

- Properties of expected value
  - Given any random variables X, Y and constants a, b: $E[aX + bY] = aE[X] + bE[Y]$
  - Suppose X and Y are independent, then $E[XY] = E[X]E[Y]$

Variance and Covariance

- Vocariance of two random variables X, Y is defined as:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

- by definition, Var[X]=Cov(X, X)
- Correlation coefficient: the correlation coefficient of X, Y is defined as:

$$\rho = \frac{Cov(X, Y)}{\sqrt{Var[X]}\sqrt{Var[Y]}}$$

- Properties of Covariance

- Given any random variables X, Y, X and any constants a, b $\in \mathbb{R}$
- Cov(X, Y)=Cov(Y, X)
- Cov(X, Y+Z)=Cov(X, Y)+Cov(X, Z)
- Cov(X, a)=0
- Cov(aX, bY)=abCov(X, Y)
- Cov(X, Y)=0 if X and Y are independent

- Properties of Variance
  - Var[X+Y]=Var[X]+Var[Y]+2Cov(X, Y)
  - If X and Y are independent, then: $Var[X+Y] = Var[X] + Var[Y]$
  - $Var[\Sigma_{i=1}^n X_i = \Sigma_{i=1}^n Var[X_i] + 2\Sigma_{1\le i\le j\le n} Cov(X_i, X_j)$
  - if $X_1, ..., X_n$ are independent, then:

$$Var[\Sigma_{i=1}^n X_i] = \Sigma_{i=1}^n Var[X_i]$$

Marginal Distributions

- marginal distribution: given a random vector (X, Y), the distribution of X or Y alone is said to be the marginal distribution of X or Y
- discrete random vector: marginal probability mass function of X is given by $P(X = x_i) = \Sigma_j P(X = x_i, Y = y_j) = \Sigma_j p_{ij}$
- continuous random vector: marginal density function of X is given by $f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$
- general case: marginal cdf of X is given by $F_X(x) = F(x, \infty)$
- $E[h(X)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)f(x, y)dxdy$

Conditional Distributions

- $$P(X \in I|B) = \frac{P(\{X \in I\} \cap B)}{P(B)}$$
- (X, Y) discrete random vector: conditional probability mass function of X given $Y = y_j$ is given by: for any x_i:

$$P(X = x_i|Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{\text{joint}}{\text{marginal}}$$

- (X, Y) continuous random vector: conditional density function of X given by Y=y is defined by for any x_i:

$$f(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{\text{joint}}{\text{marginal}}$$

Conditional distribution and Independence

- if X and Y are independent random variables, then the conditional distribution of X given Y=y coincides with the marginal distribution of X for any y
- discrete random vector:

$$P(X = x_i|Y = y_i) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = P(X = x_i)$$

- Continuous random vector:

$$f(x|y) = \frac{f(x, y)}{f_Y(y)} = f_X(x)$$

Conditional expectation:

- conditional expectation definition: given a random vector (X, Y) and a function h: $\mathbb{R} \to \mathbb{R}$
- discrete random vector: $E[h(Y)|X = x] = \Sigma_j h(y_j)P(Y = y_j|X = x)$
- continuous random vector: $E[h(Y)|X = x] = \int_{-\infty}^{\infty} h(y)f(y|x)dy$
- Properties:
  - given any random variables X, Y, Z and any constants a and b, and any function $h : \mathbb{R} \to \mathbb{R}$, we have:
  - $E[aY + bZ|X = x] = aE[Y|X = x] + bE[Z|X = x]$
  - $E[h(X)Y|X = x] = h(x)E[Y|X = x]$
  - if X and Y are independent, then $E[Y|X = x] = E[Y]$ for any x

Conditional Expectation E[Y—X] as a random variable

- definition: E[Y—X]=h(X) where h(x)=E[Y—X=x]
- E[Y—X] is a random variable and a function of X; X can be regarded as a constant when visualizing E[Y—X]

Tower Property

- Conditional Expectation: E[E[Y—X]]=E[Y]
- Law of total probability:
  - discrete: $P(A) = \Sigma_i P(A|B_i)P(B_i)$
  - let Y be a continuous random variable with density $f_Y(y)$, then: $P(A) = \int_{-\infty}^{\infty} P(A|Y = y)f_Y(y)dy$

Multivariate Normal Distributions

- Definition of $N(\mu, \Sigma)$: random vector $X = (X_1, ..., X_d)'$ takes values in $\mathbb{R}^d$ with joint density:

$$f(x) = \frac{1}{\sqrt{(2\pi)^d|\Sigma|}} exp\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\}$$

- $\mu = (\mu_1, ..., \mu_d)'$ is a mean vector; $E[X_i] = \mu_i$ for every i
- $\Sigma = [\Sigma_{ij}]$ is the covariance matrix; it is a symmetric positive definite dxd matrix and Cov$(X_i, X_j) = \Sigma_{ij}$
- standard multivariate normal distribution: $N(0, I_d)$, I_d denotes the identity dxd matrix