Name: ________________________ **Viewing Location:** ____________________

EE 464 Final Exam

May 7, 2003

Inst: C.W. Walker

Problem	Points	Score	Problem	Points	Score
1	4		11	4	
2	4		12	4	
3	4		13	4	
4	4		14	4	
5	4		15	4	
6	4		16	4	
7	4		17	4	
8	4		18	12	
9	4		19	9	
10	4		20	11	
Total	=>	=>	=>	100	

Instructions and Information:

1) Print your name and location at the top of the page.
2) Make sure your exam has 20 problems and 25 numbered pages. Pages 8-9 are for extra workspace.
3) This is a closed book exam. A formula sheet is provided at the end of the exam. You may use a calculator. **You have 2 hours to take this exam**.
4) Partial credit will be given but you must **show your work (when appropriate) to receive any credit**.
5) **Circle or box your final answers**.

Problem 1. Let $\Omega = \{u,\ v,\ w\}$. Find the subset of Ω that one would need to add (if any) to the following list of subsets in order that the list constitutes a sigma field.

$$\phi,\ \{u,\ v\},\ \{u,\ w\},\ \{v,\ w\},\ \{u\},\ \{v\},\ \{w\}.$$

a. $\{u,\ v,\ w\}$ b. {uv, w} c. {u, vw} d. {uv, vw} e. None of these.

Problem 2. An urn contains 3 red and 4 white balls. Three balls are chosen at random. Compute the probability of choosing 2 red balls if the sampling is done without replacement.

a. 0.6 b. 0.3 c. 0.5 d. None of these.

Problem 3. Compute the probability of a pair in poker (exactly 2 cards of equal face value and 3 cards of different face values).

a. $\frac{\binom{13}{1}\binom{4}{2}\binom{48}{3}}{\binom{52}{5}}$ b. $\frac{\binom{13}{1}\binom{4}{2}\binom{12}{3}4^3}{\binom{52}{5}}$ c. $\frac{\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{3}}{\binom{52}{5}}$ d. $\frac{\binom{52}{2}\binom{50}{3}}{\binom{52}{5}}$ e. None of these.

Problem 4. If $P(A|B) > P(A|C)$ then it follows that $P(B) > P(C)$.

a. True b. False c. Cannot say based on information given.

Problem 5. Suppose that in a certain test to detect a disease it is known that the probability that a person tests positive given that the person has the disease is 0.99, and the probability that a person tests positive given that the person does not have the disease is 0.005 and the probability that a randomly selected person has the disease is 0.001. Find the probability that a person has the disease given that the person tests positive for the disease.

a. 0.015 b. 0.165 c. 0.325 d. None of these.

Problem 6. The joint *pdf* for (X, Y) is given by

$$f_{XY}(x, y) = \begin{cases} k, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the constant k that makes this a valid *pdf*.

a. 2 b. 1 c. $\sqrt{2}$ d. None of these.

Problem 7. The joint *pdf* for (X, Y) is given by

$$f_{XY}(x,y) = \begin{cases} kxy, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the constant k that makes this a valid *pdf*.

a. 1 b. $\sqrt{2}$ c. 8 d. None of these.

Problem 8. You get to appear on a game show. There are 3 doors (numbered 1 to 3)with prizes behind them. One door has a car as a grand prize, another door has a television and another door has a goat as a prize. You cannot see what is behind the doors. The game show host asks you to pick a door. Say you pick door number 2. Without revealing what is behind the door you chose, the host then opens door number 3 and reveals that it has the goat as a prize (note the host knows what is behind all the doors and always opens a door that does not have the grand prize behind it). At this point the host offers to let you exchange your door for door number 1. To maximize your chances of winning the car what should you do?

a. Keep door number 2. b. Exchange your door for door number 1.

c. It does not matter since the probability of winning now is 0.5..

Problem 9. The continuous random variable X has *pdf*

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

Find $P(X \leq 2)$.

a. e^{-4} b. $1 - \mathrm{e}^{-4}$ c. e^{-2} d. $1 - \mathrm{e}^{-2}$ e. None of these.

Problem 10. Suppose $X \sim U(0,\ a)$ (uniform) and $Y \sim U(0,\ a)$. Let $Z = \max\{X, Y\}$. Find a so that $f_Z(z) = F_X(z)$.

a. $a = 1/2$ b. $\mathrm{a} = 2$ c. $\mathrm{a} = 4$ d. $\mathrm{a} = \sqrt{2}$ e. None of these.

Problem 11. The continuous random variable X has *pdf*

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

Let $Y = e^{-X}$. Then the *pdf* of Y is which of the following.

a. exponential   b. uniform   c. logarithmic   d. None of these.

Problem 12. Suppose the random variable X has mean 2 and variance 6.
Let $Y = X^2 + 1$. Find the mean of Y.

a. 5 b. 9 c. 11 d. Cannot say based on information given. e. None of these.

Problem 13. Suppose the random variable X has mean 2 and variance 6.
Let $Y = X^2 + 1$. Find the variance of Y.

a. 39 b. 22 c. 36 d. Cannot say based on information given. e. None of these.

Problem 14. Suppose $X \sim N(2,\ 5)$,i.e., X is normally distributed with mean 2 and variance 5. Find the value of b that minimzes $E\left[(X-b)^2\right]$.

a. 5   b. 2   c. 4   d. 1   e. None of these.

Problem 15. Suppose $X \sim N(2,\ 5)$. Find $E\left[X^4\right]$.

a. 75   b. 91   c. 25   d. 211   e. None of these.

Problem 16. Suppose the random variable X has mean 0 and variance 2.
Let $Y = 2X + 1$. Find the correlation coefficient, r_{XY}.

a. 1 b. 2 c. $\sqrt{2}$ d. Cannot say based on information given. e. None of these.

Problem 17. Suppose $X \sim N(2,\ 4)$, and $Y \sim N(0,\ 6)$,. Let $Z = X + Y$.
Find the variance of Z.

a. 24 b. 10 c. 6 d. Cannot say based on information given. e. None of these.

Problem 18. The two-dimensional continuous random variable (X, Y) has joint *pdf*

$$f_{XY}(x,y) = \begin{cases} x+y, & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

a. Compute $f_{Y|X}(y|x)$.

b. Find $E[Y|X = x]$.

Problem 19. Suppose Z_1 and Z_2 are bernoulli random variables with

$$E(Z_1) = E(Z_2) = p, \text{ and } Var(Z_1) = Var(Z_2) = pq.$$

Let $X = Z_1 + Z_2$ and $Y = Z_1 Z_2$. Find

a. $E(Y|X = x)$.

b. $E[E(Y|X)]$.

c. $Var[E(Y|X)]$.

d. $Var(Y|X = x)$.

e. $E[Var(Y|X)]$.

f. Best MSE predictor of Y given $X = x$.

g. Best linear MSE predictor of Y.

Extra workspace for Problem 19.

Problem 20. Suppose we toss a fair coin 50 times. Let X denote the number of heads obtained.

a. Find $P(22 < X \leq 28)$ exactly.

b. Find $P(22 < X \leq 28)$ using the normal approximation. Write your answer using

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du.$$

Extra workspace. If you use this space for work to be graded reference it from the given problem.

Extra workspace. If you use this space for work to be graded reference it from the given problem.

Final Exam Notes

Definition: The *sample space* Ω is the set of all possible outcomes of a random experiment.

Definition: An *event* is a (particular kind of) subset of the sample space.

Rules for Events

i. $\emptyset$ and Ω are events.

ii. If $\mathbf{A}$ and $\mathbf{B}$ are events then so are $\mathbf{A} \cap \mathbf{B}$, $\mathbf{A} \cup \mathbf{B}$, $\mathbf{B} \backslash \mathbf{A}$ and $\overline{\mathbf{A}} = \Omega \backslash \mathbf{A}$.

iii. If $A_1,\ A_2, \ldots,$ are events then so are

$$A_1 \cup A_2 \cup \cdots = \bigcup_{n=1}^{\infty} \mathbf{A}_n \quad \text{and} \quad A_1 \cap A_2 \cap \cdots = \bigcap_{n=1}^{\infty} \mathbf{A}_n.$$

Definition: Let $\mathbf{X}$ be a nonempty set. An *algebra* of sets on $\mathbf{X}$ is a nonempty collection $A \in P(\mathbf{X})$ which is closed under finite unions and complements, i.e.,

i. if $E_1,\ E_2, \ldots,\ E_n \in A$ then $\bigcup_{k=1}^{n} E_k \in A$.

ii. if $E \in A$ then $\overline{E} \in A$.

Definition: A σ-*algebra* (or σ-*field*) is an algebra which is closed under countable unions. So, a collection F of subsets of $\mathbf{X}$ is called a σ-field if

i. $\emptyset \in F$.

ii. $A_1,\ A_2, \ldots \in F \Rightarrow \bigcup_{k=1}^{\infty} A_k \in F$.

iii. $A \in F \Rightarrow \overline{A} \in F$.

Definition: ${}_nP_n$ is the total number of ways of arranging or permuting n different objects. Note:

$${}_nP_n = n!.$$

Definition: ${}_nP_r$ is the total number of ways of permuting r of n objects. Note:

$${}_nP_r = \frac{n!}{(n-r)!}.$$

Definition: C is the number of ways of choosing r of n objects disregarding order.

$$C = \binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

The *binomial theorem*:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Definition: The *conditional probability* of event B occurring given that event A occurs is defined to be

$$P(B|A) = \begin{cases} \dfrac{P(B \cap A)}{P(A)}, & P(A) > 0 \\ 0, & P(A) = 0. \end{cases}$$

Another way of writing this is

$$P(B|A) = \begin{cases} \dfrac{P(A|B)P(B)}{P(A)}, & P(A) > 0 \\ 0, & P(A) = 0. \end{cases}$$

Let Ω be a sample space and assume $P(A) > 0$. Given a partition $B_1,\ B_2, \ldots, B_k$ of the sample space Ω, i.e., events $B_1,\ B_2, \ldots, B_k$ such that $P(B_i) > 0$, $i = 1, 2, \ldots, k$, $B_i \cap B_j = \emptyset$, $i \neq j$ and $\bigcup_{i=1}^{k} B_i = \Omega$ then (law of total probability)

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)$$

and (Bayes' Rule)

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)}.$$

A Sum Formula:

$$\sum_{k=M}^{\infty} a^k = \frac{a^M}{1-a}, \quad 0 < a < 1.$$

Definition: The *distribution function* of a random variable X is the function

$$F_X = F : \mathbf{R} \to [0,1]$$

given by

$$F(x) = P(\{\omega : X(\omega) \leq x\}) \ \forall x \in \mathbf{R}.$$

Notation: $\{\omega : X(\omega) \leq x\}$ is abbreviated $\{X \leq x\}$. So we write

$$F(x) = P(X \leq x) \text{ or } F_X(x) = P(X \leq x).$$

F_X continuous implies $P(X = x) = 0 \ \forall x$ and

$$P(a < X \leq b) = \int_a^b f(u)du.$$

Definition: If X is continuous then

$$f_X(x) = \frac{dF_X(x)}{dx}$$

is called the *probability density function* of X. We may write $f(x) = f_X(x)$ when dealing with just one random variable.

Definition: If X is discrete and $P(X = x_i) = p_i$ then

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

is called the *probability mass function*.

Here, $f(x_i) = p_i$ and

$$\delta(u) = \begin{cases} 1, & u = 0 \\ 0, & elsewhere. \end{cases}$$

In the continuous case

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

and in the discrete case

$$F_X(x) = \sum_{u:u\leq x} f_X(u).$$

<u>Normal or Gaussian:</u> $X \sim N(\mu, \sigma^2)$. It has the form

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbf{R}.$$

The corresponding distribution function is

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma}\ e^{-(y-\mu)^2/2\sigma^2} dy.$$

<u>Uniform:</u> $X \sim U(x_1, x_2)$. It has density

$$f(x) = \begin{cases} \frac{1}{x_2 - x_1}, & x_1 \leq x < x_2 \\ 0, & elsewhere. \end{cases}$$

<u>Bernoulli:</u> X is Bernoulli if it can take on only two values.

<u>Binomial:</u> $X \sim B(n, p)$.. X is binomial of order n if

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \ k = 0, 1, 2 \ldots, n$$

where $q = 1 - p$. The mean of X is np. We can think of the binomial distribution in terms of a sum of Bernoulli trials.

<u>Geometric:</u> X is geometric if

$$P(X = k) = (1 - p)^{k-1} p.$$

Here X can be thought of as the number of trials needed before some event A occurs for the first time.

<u>Definition:</u> The *conditional distribution* $F(x|M)$ of the random variable X, given the event M occurs, is defined as

$$F(x|M) = P(X \leq x|M) = \frac{P(X \leq x, M)}{P(M)}, \quad P(M) \neq 0.$$

Definition: The *conditional density* $f(x|M)$ is the derivative of $F(x|M)$ with respect to x, i.e.,

$$f(x|M) = \frac{dF(x|M)}{dx} = \lim_{\Delta x \to 0} \frac{P(x < X \le x + \Delta x|M)}{\Delta x}.$$

Special Cases

i. Let M be the event $\{X \le a\}$ where $P(X \le a) \ne 0$. Then,

$$f(x|X \le a) = \begin{cases} 0, & x \ge a \\ \dfrac{f(x)}{F(a)}, & x < a. \end{cases}$$

ii. Let M be the event $\{b < X \le a\}$ where $F(a) \ne F(b)$. Then,

$$f(x|b < X \le a) = \begin{cases} \dfrac{f(x)}{F(a) - F(b)}, & b \le x < a \\ 0, & else. \end{cases}$$

Let $A_1, \ldots, A_n$ be a partition of Ω. Let $B = \{X \le x\}$. Then

$$P(A|X \le x) = \frac{F(x|A)}{F(x)} P(A).$$

$$P(A|x_1 < X \le x_2) = \frac{F(x_2|A) - F(x_1|A)}{F(x_2) - F(x_1)} P(A).$$

$$P(A) = \int_{-\infty}^{\infty} P(A|X = x) f(x) dx.$$

The above expression is the continuous version of the total probability theorem.

$$f(x|A) = \frac{P(A|X = x) f(x)}{\int_{-\infty}^{\infty} P(A|X = x) f(x) dx}.$$

The above expression is the continuous version of Bayes' theorem.

A Sum Formula:

$$\sum_{k=M}^{N} a^k = \frac{a^M - a^{(N+1)}}{1 - a}, \; a \ne 1.$$

9.1 Finding the Distribution of $g(X)$

9.1.2 Continuous Case

Here X is a continuous random variable and g is a continuous function. So, $Y = g(X)$ is a continuous random variable. We seek the *pdf* of Y, i.e., $f_Y(y)$.

General Procedure:

i. Obtain $F_Y(y) = P(Y \leq y)$ by finding the event A (in the range space of X) which is equivalent to the event $\{Y \leq y\}$.

ii. Differentiate $F_Y(y)$ to get $f_Y(y)$.

iii. Determine those values in the range space of Y for which $f_Y(y) > 0$.

Theorem: Let X be a continuous random variable with *pdf* $f_X(x) > 0$ for $a < x < b$. Suppose that $y = g(x)$ is a strictly monotone (strictly increasing or strictly decreasing) function of x. Assume that g is differentiable (and hence continuous) for all x. Then $Y = g(X)$ has *pdf*

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

where x is expressed in terms of y, i.e., $x = g^{-1}(y)$. Hence.

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right|.$$

If g is increasing then g is nonzero for those values of y satisfying $g(a) < y < g(b)$. If g is decreasing then g is nonzero for y satisfying $g(b) < y < g(a)$.

Theorem: Let X be a continuous random variable with *pdf* $f_X(x)$. Let $Y = X^2$. Then the *pdf* of Y is

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right].$$

9.2 Expectations

9.2.1 Discrete Case

Definition: The *expected value* of X is given by

$$E(X) = \sum_i x_i f(x_i)$$

whenever this sum exists.

Lemma: $E[g(X)] = \sum_i g(x_i) f_X(x_i)$ where $f_X(x_i) = P(X = x_i) = p_i$.

9.2.2 Continuous Case

Definition: The *expected value* of X is given by

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

whenever this integral exists.

Lemma: $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$.

9.3 Variance

9.3.1 Discrete Case

Definition: The *variance* of X is given by

$$Var(X) = \sum_i (x_i - \mu)^2 f_X(x_i)$$

where $\mu = E(X)$. Thus, $Var(X) = E\left[(X - \mu)^2\right]$.

9.3.2 Continuous Case

Definition: The *variance* of X is given by

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx$$

where $\mu = E(X)$. Again, $Var(X) = E[(X-\mu)^2]$.

9.4 Examples and Additional Results

Theorem: Let X be binomially distributed with parameters n, p (write $X \sim B(n,p)$). Then $E(X) = np$.

Theorem: $Var(X) = E(X^2) - [E(X)]^2$.

An Interpretation of Expectation:

$$\min_b E\left[(X-b)^2\right] = E\left[(X-E(X))^2\right].$$

9.5 Moments

Definitions: For $k = 1, 2, 3, \ldots$, the *kth moment* of X is

$$m_k = E\left[X^k\right]$$

and the *kth central moment* of X is

$$\mu_k = E\left[(X-E(X))^k\right].$$

Note: The 2nd central moment of X is the variance of X, $Var(X) = \sigma_X^2$. The *standard deviation* of X is $\sigma_X = \sqrt{\sigma_X^2}$.

Normal Case: Consider the mean-zero normal density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}.$$

Claim: For $n \geq 1$,

$$E[X^n] = \begin{cases} 0, & n \text{ odd}, \\ 1 \cdot 3 \cdot 5 \cdots (n-1)\sigma^n, & n \text{ even}. \end{cases}$$

10.1 Joint Distribution and Density

Definition: The *joint distribution function*, $F_{XY}(x,y)$, of two random variables, X, Y, is defined by

$$F_{XY}(x,y) = P(X \leq x, Y \leq y).$$

Definition: The *joint density function*, $f_{XY}(x,y)$, of two random variables, X, Y, is defined by

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}.$$

Therefore,

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v) du dv.$$

Definition: In the discrete case we have the *joint probabilities*

$$p_{ik} = P(X = x_i, Y = y_k).$$

Definitions: For two random variables X and Y, $F_X(x)$ is called the *marginal distribution* of X and $f_x(x)$ is called the *marginal density* of X. Similarly, for $F_Y(y)$ and $f_Y(y)$.

Claim:

$$F_X(x) = F(x,\infty),\ F_Y(y) = F(\infty, y),$$
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy,\ f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

Definition: In the discrete case we have *marginal probabilities*, p_i and q_k, where $p_i = P(X = x_i)$ and $q_k = P(Y = y_k)$.

Here

$$p_i = \sum_k p_{ik}, \quad q_k = \sum_i p_{ik}.$$

Definition: We say (X, Y) is *uniformly distributed* over a region R in the Euclideam plane if

$$f(x,y) = \begin{cases} \frac{1}{\text{area of R}}, & \text{for } (x,y) \in R, \\ 0, & \text{elsewhere.} \end{cases}$$

10.3 One Function of Two Random Variables

Given random variables X and Y and a function $g(x,y)$, we form the random variable

$$Z = g(X,Y).$$

We find the distribution function as

$$F_Z(z) = P(Z \leq z) = P(g(X,Y) \leq z) = P((x,y) \in D_z)$$

where, D_z is the region in the xy-plane consisting of the set

$$\{(x,y) : x \in R_X,\ y \in R_Y,\ g(x,y) \leq z\}.$$

So,

$$F_Z(z) = \int\int_{D_z} f(x,y)dxdy$$

where, $f(x,y)$ is the joint density of X and Y.

10.4 Two Functions of Two Random Variables

Here we have two random variables X and Y and form

$$Z = g(X,Y),\ W = h(X < Y).$$

$$F_{ZW}(z,w) = P(Z \leq z, W \leq w) = P(g(X,Y) \leq z, h(X,Y) \leq w).$$

Theorem: Let X and Y be two random variables and let $Z = g(X,Y)$ and $W = h(X,Y)$. Let $g(x,y) = z$ and $h(x,y) = w$ have solutions $(x_n, y_n), n = 1, 2, \ldots$. Then

$$f_{ZW}(z,w) = \frac{f_{XY}(x_1,y_1)}{|J(x_1,y_1)|} + \frac{f_{XY}(x_2,y_2)}{|J(x_2,y_2)|} + \cdots$$

where J is the Jacobian defined as

$$J(x,y) = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}^{-1}$$

Theorem: Let (X,Y) be a two-dimensional random variable and assume X and Y are independent. Let $W = XY$. Then

$$f_W(w) = \int_{-\infty}^{\infty} f_X(z) f_Y(w/z) \left|\frac{1}{z}\right| dz.$$

Theorem: Let (X,Y) be a two-dimensional random variable and assume X and Y are independent. Let $Z = X + Y$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(w) f_Y(z - w) dw.$$

Theorem: Let X and Y be two independent random variables each of which may assume only nonnegative integral values. Let $p_k = P(X = k), k = 0, 1, 2, \ldots$. Let $q_r = P(Y = r), r = 0, 1, 2, \ldots$. Let $W = X + Y$ and let $w_i = P(W = i)$. Then

$$w_i = \sum_{k=0}^{i} p_k q_{i-k}, \;\; i = 0, 1, 2 \ldots.$$

Given random variables X and Y, let $Z = g(X, Y)$. The expected value of Z is given by

$$E(Z) = \int_{-\infty}^{\infty} z f_Z(z) dz.$$

Theorem:

$$E(Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy.$$

If X and Y are discrete, then

$$E(Z) = E(g(X, Y)) = \sum_{i,k} g(x_i, y_k) p_{ik}.$$

Linearity: As in the 1-dim case, the 2-dim expectation operator is linear. In particular,

$$E(X + Y) = E(X) + E(Y).$$

Definition: The *covariance* of two random variables X and Y is

$$C_{XY} = E\left[(X - \mu_X)(Y - \mu_Y)\right],$$

where $\mu_x = E(X), \;\mu_Y = E(Y)$.

Note: $C_{XY} = E(XY) - E(X)E(Y)$.

Definition: The *correlation coefficient*, r_{XY}, of random variables X and Y is

$$r_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}, \;\; \sigma_X \sigma_y \neq 0.$$

Note: We define $C_{XY} = 0$ if $\sigma_X = 0$ or $\sigma_Y = 0$.

Claim: $|r_{XY}| \leq 1$ (i.e., $|C_{XY}| \leq \sigma_X \sigma_Y$).

Theorem: $r_{XY}^2 = 1$ if and only if $Y = aX + b$ for some constants a, b.

Definition: Two random variables X and Y are called *uncorrelated* if their covariance is zero, i.e., $C_{XY} = 0$.

Theorem: If X and Y are independent then they are uncorrelated.

Variance of $X + Y$

Consider $Z = X + Y$. Then $E[Z] = \mu_Z = \mu_X + \mu_Y$.

$$\sigma_Z^2 = \sigma_X^2 + 2r_{XY}\sigma_X\sigma_Y + \sigma_Y^2.$$

If X and Y are uncorrelated then $r_{XY} = 0$ and

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2.$$

It then follows that if X and Y are independent

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2.$$

Theorem: Let $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$ and let X and Y be independent. Then $Z = X + Y$ is also normal and

$$E(Z) = \mu_X + \mu_Y, \quad Var(Z) = Var(X) + Var(Y).$$

More generally, if we have the above conditions and $Z = aX + bY$, then

$$Z \sim N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2).$$

11.3 Conditional Distributions

$$f_{XY}(x, y | x_1 < X \leq x_2) = \begin{cases} \dfrac{f_{XY}(x, y)}{F_X(x_2) - F_X(x_1)}, & x_1 < x \leq x_2, \\ 0, & \text{elsewhere.} \end{cases}$$

$$f_Y(y|X=x) = \lim_{\Delta x \to 0} f_Y(y|x < X \le x+\Delta x) = \frac{f_{XY}(x,y)}{f_X(x)}.$$

Notation:

$$f(y|x) = f_Y(y|x) = f_Y(y|X=x), \quad f(x|y) = f_X(x|y) = f_X(x|Y=y),$$

$$f(y|x) = \frac{f(x,y)}{f(x)}, \quad f(x|y) = \frac{f(x,y)}{f(y)}.$$

Notation: If X and Y are independent

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

and similarly $f_X(x|y) = f_X(x)$.

Now

$$f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_Y(y|x)f_X(x)}{f_Y(y)}.$$

Also

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

and

$$f_{XY}(x,y) = f_Y(y|x)f_X(x).$$

Hence,

$$f_Y(y) = \int_{-\infty}^{\infty} f_Y(y|x)f_X(x)dx.$$

This is total probability. We thus get Bayes' theorem for densities

$$f_X(x|y) = \frac{f_Y(y|x)f_X(x)}{\displaystyle\int_{-\infty}^{\infty} f_Y(y|x)f_X(x)dx}.$$

Discrete Type: Let $p_i = P(X = x_i),\ p_{ik} = P(X = x_i,\ Y = y_k)$. Then.

$$P(Y = y_k|X = x_i) = \frac{P(X = x_i,\ Y = y_k)}{P(X = x_i)} = \frac{p_{ik}}{p_i}.$$

11.4 Conditional Expected Values

Definition: The *conditional mean of* $g(Y)$ *given* $X \le x$ is given by

$$E\left[g(Y)|X \le x\right] = \int_{-\infty}^{\infty} g(y)f(y|X \le x)dy.$$

Definition: The *conditional mean of $g(Y)$ given $X = x$* is given by

$$E\left[g(Y)|X = x\right] = \int_{-\infty}^{\infty} g(y) f(y|x) dy.$$

In particular, we have the conditional mean of Y given $X = x$

$$\mu_{Y|X} = E\left[Y|X = x\right] = \int_{-\infty}^{\infty} y f(y|x) dy$$

and the conditional variance of Y given $X = x$

$$\sigma^2_{Y|X} = E\left[(Y - \mu_{Y|X})^2 | X = x\right] = \int_{-\infty}^{\infty} (y - \mu_{Y_X})^2 f(y|x) dy.$$

Notation: $E\left[g(Y)|x\right] = E\left[g(Y)|X = x\right]$.

Preceding developments lead to the following theorem.

Theorem:

$$E\left[g(X, Y)|M\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y|M) dx dy$$

for an event M.

$$E\left[g(X, Y)|X = x\right] = \int_{-\infty}^{\infty} g(x, y) f(y|x) dy.$$

Note that the conditional mean of Y given $X = x$ is itself a function of x:

$$E\left[Y|X = x\right] = \int_{-\infty}^{\infty} y f(y|x) dy.$$

Then $E[Y|X]$ is a random variable and

$$E\left[E(Y|X)\right] = E(Y).$$

Similarly,

$$E\left[E\left(g(X, Y)|X\right)\right] = E\left[g(X, Y)\right].$$

11.5 Mean Square Estimation

Nonlinear MS Estimation:

$E(Y|X=x)$ is the best MSE predictor of Y given $X=x$.

Linear MS Estimation:

Theorem: Suppose that $E(X^2)$ and $E(Y^2)$ are finite and X and Y are not constant. Then the best (in the MS sense) zero intercept linear predictor of Y ($\hat{Y}=a_0X$) is obtained by taking

$$a_0 = \frac{E(XY)}{E(X^2)}$$

while the best linear predictor of Y ($\hat{Y}=a_1X+b_1$) is

$$a_1 = \frac{Cov(X,Y)}{Var(X)}, \quad b_1 = E(Y) - a_1E(X).$$

12.0 Sequences of Random Variables

Definition: A *random vector* is denoted

$$\underline{X} = [X_1, X_2, \dots, X_n]$$

where each $X_i,\ i = 1, 2, \dots, n$ is a random variable.

12.9 Central Limit Theorem (CLT)

Theorem(CLT): Let $X_1, \dots, X_n$ be n independent random variables with finite means μ_i and finite non-zero variances σ_i^2. Let

$$S_n = X_1 + \cdots + X_n,$$

so

$$\mu_S = E(S_n) = \sum_{i=1}^{n} \mu_i, \qquad \sigma_S^2 = Var(S_n) = \sum_{i=1}^{n} \sigma_i^2.$$

Then

$$\lim_{n\to\infty} \frac{S_n - \mu_S}{\sigma_S} \sim N(0,\ 1).$$

Let us consider $\overline{X} = S_n/n$. Then

$$\lim_{n\to\infty} \frac{\overline{X} - E(\overline{X})}{\sqrt{Var(\overline{X})}} \sim N(0,\ 1).$$

Restricted Case

Theorem(CLT): Let $f(\cdot)$ be a density with mean μ and variance σ^2. Let $\overline{X}_n$ be the sample mean of n independent random samples of size n from $f(\cdot)$. Let

$$Z_n = \frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}.$$

Then,

$$\lim_{n\to\infty} Z_n \sim N(0,\ 1).$$