A review of Chapter 3 and Chapter 4

STA 524, Fall 2008

Ruriko Yoshida
Dept. of Statistics University of Kentucky

www.ms.uky.edu/~ruriko

Discrete random variables

Definition A **random variable** X is a function from Ω to $\mathbb{R}$.

Definition We say X a **discrete random variable** if X can take a sequence of different values.

Definition If X is a discrete random variable and has a discrete distribution the **probability function** of X is defined as the function F s.t.

$$f(x) = Pr(X = x), \forall x \in \mathbb{R}.$$

Continuous random variables

Definition We say X a **continuous random variable** if there is a continuous nonnegative function f defined on $\mathbb{R}$ such that

$$Pr(X \in A) = \int_A f(x)dx, \forall A \subset \mathbb{R}.$$

This function f is called the **probability density function** of X.

Note For every p.d.f f of X must satisfy:

1. $f(x) \geq 0$ for all $x \in \mathbb{R}$.

2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Distribution functions

Definition The **distribution function** F of a r.v. X is a function defined for each real number $x \in \mathbb{R}$ s.t.

$$F(x) = Pr(X \leq x),\ x \in \mathbb{R}.$$

Thm The function $F(x)$ is non-decreasing as x increases.

Thm

$$\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to \infty} F(x) = 1.$$

Thm

$$F(x) = F(x^+),\ \forall x \in \mathbb{R}.$$

Distribution functions

Thm

$$Pr(X > x) = 1 - F(x).$$

Thm

$$Pr(x_1 \leq X \leq x_2) = F(x_2) - F(x_1), \ \forall x_1 \leq x_2 \in \mathbb{R}.$$

Thm

$$Pr(X < x) = F(x^-), \ \forall x \in \mathbb{R}.$$

Thm

$$Pr(X = x) = F(x) - F(x^-), \ \forall x \in \mathbb{R}.$$

Joint probability (density) functions

Definition Suppose X and Y are discrete r.v. Then the **joint probability function** of X and Y is defined as a function f s.t.

$$f(x,y) = Pr(X = x \text{ and } Y = y), \forall x, y \in \mathbb{R}.$$

Definition Suppose X and Y are continuous r.v. If there is a nonnegative continuous function f defined over $\mathbb{R}^2$ s.t.

$$Pr[(X,Y) \in A] = \int\int_A f(x,y)dxdy, \forall A \subset \mathbb{R}^2,$$

then this function $f(x,y)$ is called the **joint probability density function** of X and Y.

Joint probability density functions

Suppose $f(x, y)$ is the joint density function of X and Y. Then it must satisfy:

1. $f(x, y) \geq 0$ for all x, $y \in \mathbb{R}$.

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Joint distribution functions

Definition The **joint distribution function** of X and Y is defined as a function F over $\mathbb{R}^2$ s.t.

$$F(x,y) = Pr[X \leq x \text{ and } Y \leq y], \forall x, y, \in \mathbb{R}.$$

Note

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v) du dv \text{ and } f(x,y) = \partial^2 F(x,y)/\partial x \partial y.$$

Marginal distributions

Definition If X and Y are discrete r.v., then the **marginal probability function** of X is

$$f_1(x) = Pr(X = x) = \sum_y Pr(X = x \text{ and } Y = y) = \sum_y f(x, y).$$

The **marginal probability function** of Y is

$$f_2(y) = Pr(Y = y) = \sum_x Pr(X = x \text{ and } Y = y) = \sum_x f(x, y).$$

Marginal distributions

Definition If X and Y are continuous r.v., then the **marginal probability density function** of X is

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$

The **marginal probability density function** of Y is

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Marginal distributions

Definition If X and Y are r.v., then the **marginal distribution function** of X is

$$F_1(x) = Pr(X \leq x) = \lim_{y \to \infty} Pr(X \leq x \text{ and } Y \leq y) = \lim_{y \to \infty} F(x, y).$$

The **marginal distribution function** of Y is

$$F_2(y) = Pr(Y \leq y) = \lim_{x \to \infty} Pr(X \leq x \text{ and } Y \leq y) = \lim_{x \to \infty} F(x, y).$$

Independent

Definition We say r.v. X and Y are **independent** if

$$Pr(X \in A \text{ and } Y \in B) = Pr(X \in A)Pr(Y \in B),\ \forall A, B \subset \mathbb{R}.$$

Note X and Y are independent iff

$$\begin{gathered} Pr(X \leq x,\, Y \leq y) = Pr(X \leq x)Pr(Y \leq y) \text{ iff} \\ F(x, y) = F_1(x)F_2(y) \text{ iff} \\ f(x, y) = f_1(x)f_2(y). \end{gathered}$$

Independent

Definition We say n r.v. $X_1, \cdots, X_n$ are **independent** iff $\forall A_1, \cdots, A_n \subset \mathbb{R}$,

$$Pr(X_1 \in A_1, \cdots X_n \in A_n) = Pr(X_1 \in A_1) \cdots Pr(X_n \in A_n).$$

Note $X_1, \cdots, X_n$ are independent iff

$$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n) \text{ iff}$$
$$F(x_1, \cdots, x_n) = F_1(x_1) \cdots F_n(x_n).$$

Expectations

Definition If a r.v. X has a discrete distribution,

$$\mathbb{E}(X) = \sum_{w \in \Omega} X(w) Pr(\{x\}) = \sum_{x} x f(x).$$

Definition If a r.v. X has a continuous distribution,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Expectations

Note We say $\mathbb{E}(X)$ exists iff $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$.

Thm For a nonnegative continuous r.v. $X \geq 0$,

$$\mathbb{E}(x) = \int_0^{\infty} (1 - F(x)) dx.$$

Thm If $Y = aX + b$, where $a, b \in \mathbb{R}$ are constant, then

$$\mathbb{E}(Y) = a\mathbb{E}(X) + b.$$

Thm If there is a constant $a \in \mathbb{R}$ s.t. $Pr(X \geq a) = 1$, then $\mathbb{E}(X) \geq a$. If there is a constant $b \in \mathbb{R}$ s.t. $Pr(X \leq b) = 1$, then $\mathbb{E}(X) \leq b$.

Expectations

Thm If $X_1, \cdots, X_n$ are r.v. s.t. $\mathbb{E}(X_i)$ exists for all i,

$$\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n).$$

Thm If $X_1, \cdots, X_n$ are independent r.v. s.t. $\mathbb{E}(X_i)$ exists for all i,

$$\mathbb{E}(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} \mathbb{E}(X_i).$$

Thm If $X_1, \cdots, X_n$ form a random sample with mean μ and if $Y = X_1 + \cdots + X_n$ and $M = \frac{1}{n}(X_1 + \cdots + X_n)$. Then

$$\mathbb{E}(Y) = n\mu \text{ and } \mathbb{E}(M) = \mu.$$

Variance

Definition For a r.v. X with a finite mean, the **variance** is defined to be

$$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2].$$

The **standard deviation** of X is $\sqrt{Var(X)}$.

Thm $Var(X) = 0$ iff there is a constant c s.t. $Pr(X = c) = 1$.

Thm For constants $a, b \in \mathbb{R}$, $Var(aX + b) = a^2 Var(X)$.

Thm

$$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

Thm If $X_1, \cdots, X_n$ are independent r.v.,

$$Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n).$$

Moments

Definition The expectation $\mathbb{E}(X^k)$ for a positive integer k is called the k**th moment** of X.

Note We say the kth moment exists iff $\mathbb{E}(|X|^k) < \infty$.

Thm If $\mathbb{E}(|X|^k) < \infty$ for some k then $\mathbb{E}(|X|^j) < \infty$ for all positive integer $j < k$.

Definition The expectation $\mathbb{E}[(X - \mu)^k]$ for a positive integer k is called the k**th central moment** of X.

Moment generating functions

Definition The **moment generating function** $\psi(t)$ is defined

$$\psi(t) = \mathbb{E}(\exp(tX)).$$

Note If there is a derivative ψ' around $t = 0$, then

$$\psi'(0) = \mathbb{E}(X).$$

If there is a derivative ψ^k around $t = 0$, then

$$\psi^k(0) = \mathbb{E}(X^k).$$

Moment generating functions

Thm Let X be a r.v. with the mgf ψ_1 and let $Y = aX + b$ where $a, b \in \mathbb{R}$ are constant and ψ_2 is the mgf of Y. Then

$$\psi_2(t) = \exp(bt)\psi_1(at).$$

Thm If $X_1, \cdots, X_n$ are independent r.v. with the mgf $\psi_1, \cdots, \psi_n$, respectively, then

$$\psi(t) = \prod_{i=1}^{n} \psi_i(t), \forall t \text{ with } \psi_i(t) \text{ exit,}$$

where ψ is the mgf of $X_1 + \cdots + X_n$.

Covariance and correlations

Definition The **covariance** of X and Y is defined

$$Cov(X,Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

Definition If $0 < \sigma_X^2 < \infty$ and $0 < \sigma_Y^2 < \infty$, then the **correlation** of X and Y is defined

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}.$$

Covariance and correlations

Thm (Schwarz Inequality)

$$[\mathbb{E}(XY)]^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2).$$

Thm If $\sigma_X^2 < \infty$ and $\sigma_Y^2 < \infty$, then

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Thm If X and Y are independent and $0 < \sigma_X^2 < \infty$ and $0 < \sigma_Y^2 < \infty$,

$$Cov(X, Y) = \rho(X, Y) = 0.$$

Covariance and correlations

Thm Let X be a r.v with $0 < \sigma_X^2 < \infty$ and let $Y = aX + b$ where $a, b, \in \mathbb{R}$ are constant. If $a > 0$, then $\rho(X, Y) = 1$ and if $a < 0$, then $\rho(X, Y) = -1$.

Thm If X and Y are r.v. s.t. $Var(X) < \infty$ and $Var(Y) < \infty$, then

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y).$$

Thm If $X_1, \cdots, X_n$ are r.v. with $Var(X_i) < \infty$, then

$$Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n) + 2\sum_{i<j} Cov(X_i, X_j).$$