Chapter 5. PROBABILITY DENSITIES (B)

March 1, 2011

1 Joint Distributions—Discrete and Continuous r.v.

Discrete Variables

- For two discrete variables X_1 and X_2, we write the probability that X_1 will take the value x_1 and X_2 will take the value x_2, as

$$P(\{X_1 = x_1\} \cap \{X_2 = x_2\}) = P(X_1 = x_1, X_2 = x_2).$$

which is a function of x_1 and x_2, denoted by $f(x_1, x_2)$. We call the function $f(x_1, x_2)$ **joint probability distribution** of X_1 and X_2.

- The distribution of X_1 (or X_2) alone is called the **marginal distribution**

$$f_1(x_1) = P(X_1 = x_1) = \sum_{all\ \ x_2} f(x_1, x_2)$$

and

$$f_2(x_2) = P(X_2 = x_2) = \sum_{all\ \ x_1} f(x_1, x_2)$$

- **Conditional probability distribution** of X_1 given $X_2 = x_2$ is

$$f_1(x_1|x_2) = P(X_1 = x_1|X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = \frac{f(x_1, x_2)}{f_2(x_2)}$$

for all x_2 provided $f_2(x_2) \neq 0$.

- By the definition of conditional probability distribution

$$f(x_1, x_2) = f_2(x_2) f_1(X_1 = x_1 | X_2 = x_2)$$

- **Independence**. If $f_1(x_1|x_2) = f(x_1)$ for all x_1 and x_2, then we say that X_1 and X_2 are independent. It is equivalent to

$$f(x_1, x_2) = f_1(x_1) f_2(x_2)$$

for all x_1 and x_2.

Example Let X_1 and X_2 have the joint probability distribution in the table below.

$f(x_1, x_2)$

		x_1		
		0	1	2
x_2	0	0.1	0.4	0.1
	1	0.2	0.2	0

- Find $P(X_1 + X_2 > 1)$

- find the marginal distributions $f_1(x_1)$ and $f_2(x_2)$

$f_1(x_1)$

x_1	0	1	2
$f_1(x_1)$	0.3	0.6	0.1

$f_2(x_2)$

x_2	0	1
$f_2(x_2)$	0.6	0.4

- are X_1 and X_2 independent

For n discrete variables $X_1, ..., X_n$

- Define $P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$, denoted by $f(x_1, x_2, ..., x_n)$, as the **joint probability distribution** of these discrete random variables.

- The probability distribution $f_i(x_i)$ of the individual variable X_i is called the **marginal probability** distribution of the ith random variable

- **Independence**. If

$$f(x_1, x_2, ..., x_n) = f_1(x_1)f_2(x_2)...f_n(x_n)$$

for all x_1, $x_2, ..., x_n$, we say random variables $X_1, ..., X_n$ are **independent**.

Continuous variables

- If $X_1, X_2, ..., X_k$ are k continuous random variables, we shall refer to $f(x_1, x_2, ..., x_k)$ as the **joint probability density function** or **probability density function (pdf)** of these random variables. If the probability that $a_1 \le X_1 \le b_1, a_2 \le X_2 \le b_2, ..., a_k \le X_k \le b_k$ is given by the multiple integral

$$P(a_1 \le X_1 \le b_1, a_2 \le X_2 \le b_2, ..., a_k \le X_k \le b_k) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} ... \int_{a_k}^{b_k} f(x_1, x_2, ..., x_k) dx_1 dx_2 ... dx_k$$

- Any density function $f(x_1, ..., x_n)$ must satisfy

(a) $f(x_1, ..., x_n) \ge 0$

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, x_2, ..., x_k) dx_1 dx_2 ... dx_k = 1$

- Any function satisfying (a) and (b) is a probability density function.

- joint **cumulative distribution function, CDF**

$$F(x_1, x_2, ..., x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} f(u_1, u_2, ..., u_n) du_1 du_2 ... du_n$$

- **marginal density function** for X_1

$$f_1(x_1) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(u_1, u_2, ..., u_n) du_2 ... du_n$$

for X_2

$$f_2(x_2) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(u_1, u_2, ..., u_n) du_1 du_3 ... du_n$$

...

- **marginal CDF** for X_1

$$F_k(x_k) = \int_{-\infty}^{x_k} f_k(u_k) d du_k$$

- **conditional probability density function**

$$f_1(x_1|x_2) = \frac{f(x_1, x_2)}{f(x_2)}$$

Again

$$f(x_1, x_2) = f(x_2) f_1(x_1|x_2)$$

- **independence** Variables $X_1, ..., X_n$ are independent if

$$f(x_1, ..., x_n) = f_1(x_1) f_2(x_2) ... f_n(x_n)$$

for all values $x_1, ..., x_n$, or

$$F(x_1, ..., x_n) = F_1(x_1)F_2(x_2)...F_n(x_n)$$

for all values $x_1, ..., x_n$.

- **Example** If two variables X_1 and X_2 have a joint density function

$$f(x_1, x_2) = \begin{cases} 6e^{-2x_1 - 3x_2}, & \text{if } x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

find the probabilities that

(a) the first random variable will take on a value between 1 and 2 and the second random variable a value between 2 and 3;

(b) the first random variable will take on a value less than 2 and the second random variable will take on a value greater 2.

(c) find the cumulative distribution

(d) are X_1 and X_2 independent?

solution

(a) $\int_1^2 \int_2^3 f(x_1, x_2) dx_1 dx_2 = \int_1^2 \int_2^3 6e^{-2x_1-3x_2} dx_1 dx_2 = (e^{-2} - e^{-4})(e^{-6} - e^{-9}) = 0.0003$

(b) $\int_0^2 \int_2^\infty f(x_1, x_2) dx_1 dx_2 = (e^0 - e^{-4})e^{-6} = 0.0025$

(c)

$$F(x_1, x_2) = \begin{cases} \int_0^{x_1} \int_0^{x_2} 6e^{-2u_1-3u_2} du_1 du_2, & \text{if } x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$
$$= \begin{cases} (1 - e^{-2x_1})(1 - e^{-3x_2}), & \text{if } x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

(c) Note that

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, u_2) du_2 = \begin{cases} 2e^{-2x_1}, & \text{if } x_1 > 0 \\ 0, & \text{otherwise} \end{cases}$$
$$f_2(x_2) = \int_{-\infty}^{\infty} f(u_1, x_2) du_1 = \begin{cases} 3e^{-3x_1}, & \text{if } x_1 > 0 \\ 0, & \text{otherwise} \end{cases}$$

Thus

$$f_1(x_1) f_2(x_2) = f(x_1, x_2)$$

They are independent.

2 Moments of several variables

- **Expected value (or mean, expectation) of** $g(X_1, X_2, ..., X_k)$

  In the discrete case,

$$E[g(X_1, X_2, ..., X_k)] = \sum_{x_1} \sum_{x_2} ... \sum_{x_k} g(x_1, x_2, ..., x_k) f(x_1, x_2, ..., x_k)$$

  In the continuous case,

$$\begin{aligned} E[g(X_1, X_2, ...X_n)] \; = \; & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} g(x_1, x_2, ..., x_k) \\ & \times f(x_1, x_2, ..., x_k) dx_1 dx_2 ... dx_k \end{aligned}$$

- For any random variables $X_1, ..., X_n$ and constants $a_0, a_1, ..., a_n$, we have

$$E(a_0 + a_1 X_1 + ... + a_n X_n) = a_0 + a_1 E(X_1) + ... + a_n E(X_n)$$

- **Covariance** of X_i and X_j

$$Cov(X_i, X_j) = E\{(X_i - EX_i)(X_j - EX_j)\}$$

- If X_i and X_j are independent then

$$Cov(X_i, X_j) = 0$$

- For any random variables $X_1, ..., X_n$ we have

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

- if random variables $X_1, ..., X_n$ are independent, then

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

and

$$Var(a_0 + a_1X_1 + ... + a_nX_n) = a_1^2Var(X_1) + ... + a_n^2Var(X_n)$$

- **Example** Let X have mean μ_1 and variance σ_1^2 and let X_2 have mean μ_2 and variance σ_2^2. If they are independent, find the mean and variance of

(a) $X_1 - X_2$

(b) $X_1 + X_2$

- **Example** Finding the mean and variance of $2X_1 + X_2 - 5$ if X_1 has mean 4 and variance 9 while X_2 has mean -2 and variance 5, and are independent. Find

(a) $E(2X_1 + X_2 - 5)$

$$E(2X_1 + X_2 - 5) \; = \; 2EX_1 + EX_2 - 5 = 1$$

(b) $Var(2X_1 + X_2 - 5)$

$$Var(2X_1 + X_1 - 5) = 4Var(X_1) + Var(X_2) = 41$$

- **Example** [Experimental Design] Two rods of unknown lengths a, b. A ruler can measure the length but with error having 0 mean (unbiased) and variance σ^2. Errors independent from measurement to measurement. To estimate a, b we could take separate measurements A, B of each rod.

$$E(A) = a, \qquad \mathbf{var}[A] = \sigma^2$$

$$E(B) = b, \qquad \mathbf{var}[B] = \sigma^2$$

Can we do better? YEP! Measure $a + b$ as X and $a - b$ as Y

$$\mathbf{E}[X] = a + b, \qquad \mathbf{var}[X] = \sigma^2$$

$$\mathbf{E}[Y] = a - b, \qquad \mathbf{var}[Y] = \sigma^2$$

$$\mathbf{E}\left[\frac{X+Y}{2}\right] = a, \quad \mathbf{var}\left[\frac{X+Y}{2}\right] = \frac{1}{2}\sigma^2$$
$$\mathbf{E}\left[\frac{X-Y}{2}\right] = b, \quad \mathbf{var}\left[\frac{X-Y}{2}\right] = \frac{1}{2}\sigma^2$$

So this (i.e. using $(X+Y)/2$ to estimate a, and $(X+Y)/2$ to estimate b) is better.

3 The mean and variance of the sample mean

Let the n random variables $X_1, X_2, ..., X_n$ be independent and each has the same distribution with μ and variance σ^2. The sample mean is defined as

$$\bar{X} = (X_1 + ... + X_n)/n$$

and sample variance

$$s^2 = [(X_1 - \bar{X})^2 + ... + (X_n - \bar{X})^2]/(n-1)$$

- sample mean

(a) mean:

$$\mu_{\bar{X}} = E(\bar{X}) = \mu$$

so the expected value or mean of X is the same as the mean of each observation.

(b) variance:

$$\sigma_{\bar{X}}^2 = Var(\bar{X}) = \sigma^2/n$$

so the variance of $\bar{X}$ equals the variance of a single observe divided by n.

In finance, this is a way of "risk aversion". Diversifying investments on a variety of assets can reduce risk

- sample variance

$$Es^2 = \sigma^2$$

Proof: write

$$(X_i - \bar{X})^2 = (X_i - \mu + \mu - \bar{X})^2 = (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X})$$

and thus

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n}(\mu - \bar{X})^2 + \sum_{i=1}^{n}2(X_i - \mu)(\mu - \bar{X})$$

Note that for the last term

$$\sum_{i=1}^{n}2(X_i - \mu)(\mu - \bar{X}) = -2n(\bar{X} - \mu)^2$$

Consequently

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}(X_i - \mu)^2 - n(\mu - \bar{X})^2$$

Since $E(X_i - \mu)^2 = \sigma^2$ and $E(\mu - \bar{X})^2 = \sigma^2/n$, we have

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}\sigma^2 - n\sigma^2/n = (n-1)\sigma^2.$$

Thus

$$Es^2 = E\Big[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\Big] = \sigma^2.$$