

Stat 324: Lecture 11

Bivariate Distributions

Moo K. Chung
mchung@stat.wisc.edu

March 3, 2005

1. *Motivation.* There are many practical situations where we need to deal with more than two measurement simultaneously. For instance we might be interested in the relationship between heights and weights of a population.
2. *Definition.* For discrete random variables X and Y , the *joint probability distribution* is

$$p(x, y) = P(X = x, Y = y).$$

For continuous random variables X and Y , the *joint density function* $f(x, y)$ is defined as a function satisfying

$$P((X, Y) \in A) = \int \int_A f(x, y) dx dy.$$

3. *Definition.* Marginal probability function of X is given by $p_X(x) = \sum_y p(x, y)$. This is the probability function of X . The marginal density function of X is given by $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$.
4. Random variables X and Y are independent if $p(x, y) = p_X(x)p_Y(y)$ for discrete case and $f(x, y) = f_X(x)f_Y(y)$ for continuous case for all x and y .
5. The conditional probability density of Y given $X = x$ is defined as $f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$. The conditional expectation of Y given $X = x$ is

$$\mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy.$$

6. *Problem.* If X and Y are jointly continuous random variables with joint (cumulative) distribution function $F(x, y)$, express $P(a \leq X \leq b, c \leq Y \leq d)$ using only F .
Solution. This is a fun problem testing whether you really understand the concept of a joint cumulative distribution function. Draw $\{a \leq X \leq b, c \leq Y \leq d\} \subset \mathbb{R}^2$. Then $P(a \leq X \leq b, c \leq Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$ using the fact that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. *Problem.* Suppose that the joint probability density function f of X and Y is

$$f(x, y) = \begin{cases} k(x^2 + y) & \text{for } 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

Find the value k and the marginal density functions. Are X and Y independent random variables? Justify your answer.

Solution. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = \int_0^2 \int_0^1 k(x^2 + y) dx dy = \int_0^2 k(\frac{1}{3} + y) dy = \frac{8}{3}k = 1$. So $k = \frac{3}{8}$.

$$f_X(x) = \int_0^2 \frac{3}{8}(x^2 + y) dy = \begin{cases} \frac{3}{4}(x^2 + 1) & \text{for } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$f_Y(y) = \int_0^1 \frac{3}{8}(x^2 + y) dx = \begin{cases} \frac{1}{8}(1 + 3y) & \text{for } 0 \leq y \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Since $f(0, 1) \neq f_X(0)f_Y(1)$, they are not independent.

Assigned Problems. Exercise 5.10, Exercise 5.12. Your second midterm is tentatively scheduled to be on April 5th.