4. Multivariate Random Variables

FRM Part 1: Quantitative Analysis

Devere Anthony Weaver

---

1 Jointly Distributed Random Variables

1.1 Joint PMF for Two Discrete Random Variables

Definition 1 (Joint Probability Mass Function)**.** *Let X and Y be two discrete random variables defined on sample space Ω. The **joint probability mass function** $p(x, y)$ is defined for each pair of numbers (x, y) by*

$$p(x, y) = P[X = x \cap Y = y].$$

Of course this joint PMF must also retain the properties of a univariate pmf (i.e. probabilities are greater than or equal to 0 and their sum is 1).

Definition 2 (Marginal Probability Mass Function)**.** *The **marginal probability mass functions** of X and Y, are given by*

$$p_X(x) = \sum_Y p(x, y)$$

and

$$p_Y(y) = \sum_X p(x, y).$$

1.2 Joint PDF for Two Continuous Random Variables

Definition 3 (Joint Probability Density Function)**.** *Let X, Y by two continuous random variables. Then $f(x, y)$ is the **joint probability density function** for X and Y if for any two-dimensional set A,*

$$P[(x, y) \in A] = \int\int_A f(x, y)dxdy.$$

Definition 4 (Marginal Probability Density Function)**.** *The **marginal probability density functions** of X and Y are given by*

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, \;\; -\infty < x < \infty$$

and

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

1.3 Independent Random Variables

Definition 5. *Two random variable X and Y and said to be **independent** if for every (x, y)*

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

or

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

Obviously the choice of which equation to use depends on if the random variables are discrete or continuous.

Using the definition above, this means that if we want to test if two random variables are independent, we simply need to find one (x, y) pair where this doesn't hold.

While this section only the case of two random variables, the above definitions can be extended to the general case of an n number of multiple random variables.

---

2 Expected Values, Covariance, and Correlation

Law of the Unconscious Statistician Let X and Y be jointly distributed random variables with PMF $p(x, y)$ or PDF $f(x, y)$. Then, the expected value of a function $h(x, y)$, denoted $E[h(x, y)]$, is given by

$$E[h(x,y)] = \begin{cases} \sum_X \sum_Y h(x,y) \cdot p(x,y) \\ \\ \int_{-\infty}^{x} \int_{-\infty}^{y} h(x,y) \cdot f(x,y) dx dy. \end{cases}$$

Definition 6 (Covariance)**.** *The* ***covariance*** *between two random variables X and Y is*

$$\begin{aligned} Cov(X,Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= \begin{cases} \sum_X \sum_Y (X - \mu_X)(Y - \mu_Y) p(x,y) \\ \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (X - \mu_X)(Y - \mu_Y) f(x,y) \, dx dy. \end{cases} \end{aligned}$$

Some properties of covariance include:

1. $Cov(X, Y) = Cov(Y, X)$
2. $Cov(X, X) = Var(X)$
3. $Cov(X, Y) = E(XY) - E[X]E[Y]$ (a "shortcut" formula
4. $\forall Z$ and constants a, b, c, $Cov(aX + bY + c, Z) = aCov(X, Z) + bCov(Y, Z)$.

Definition 7 (Correlation)**.** *The* ***correlation coefficient*** *of X and Y is given by*

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}.$$

Some properties of the correlation coefficient include:

1. $Corr(X, Y) = Corr(Y, X)$
2. $Corr(X, X) = 1$
3. $-1 \leq Corr(X, Y) \leq 1$
4. If a, b, c, d are all constants and $ac > 0$, then $Corr(aX + b, cY + d) = Corr(X, Y)$. In other words, we can ignore the scale factor.

One important thing to note is that if random variable X and Y are independent, then $\rho_{x,y} = 0$. However, $\rho_{x,y} = 0$ does **not** imply independence.

2.1 Moments

In the case of bivariate random variables, the mean is the first moment and it is simply the mean of the random variable's components. Let X be a continuous random variable with components (x_1, x_2). Then the first moment for X is given by

$$E[X] = [E[X_1], E[X_2]] = [\mu_1, \mu_2].$$

---

3 Conditional Distributions and Conditional Expectation

Definition 8 (Conditional Distribution)**.** *Let X, Y be two discrete random variables with joint PMF $p(x, y)$ and marginal X PMF $p_X(x)$. The, for all X such that $p_X(x) \geq 0$, the* ***conditional probability mass function of*** *Y* ***given*** *$X = x$ is given by*

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)}.$$

For the continuous case, the conditional probability density function of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}.$$

3.1 Conditional Distributions and Independence

Basically, we can use the definition of a conditional distribution to verify our idea of independence. If we evaluate the conditional distribution and end up getting as the result the marginal distribution of the conditioned variable, then they are independent. This means the marginal distribution of the other random variable carries no information about the conditioned variable.

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)} = \frac{p_X(x)p_Y(x)}{p_X(x)} = p_Y(y).$$

3.2 Conditional Expectation and Variance

Definition 9. *(Conditional Expectation) Let X, Y be two discrete random variables with conditional PMF $p_{Y|X}(y|x)$. Then the conditional expectation of Y given $X = x$ is*

$$E(Y|X = x) = \sum_{Y} y \cdot p_{Y|X}(y|x).$$

Similarly for continuous random variables,

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

These definitions can be extended to functions in the same way that univariate expectations are.