Lecture 8 - Multivariate Random Variables II

Function of random vector

- Let $g(\mathbf{X})$ be a function of a discrete random vector $\mathbf{X}$. It holds:

$$E(g(\mathbf{X})) = \sum_{\mathbf{x}} g(\mathbf{x}) P(\mathbf{X} = \mathbf{x})$$

- In particular for the case of two random variables:

$$E(g(X,Y)) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} g(x,y) P(X = x, Y = y)$$

Additivity of expected value

- For any random variables X and Y it holds:

$$E(X+Y) = EX + EY$$

- The same for the case of n random variables:

$$E(X_1 + X_2 + \ldots + X_n) = EX_1 + EX_2 + \ldots + EX_n$$

Coupon collector’s problem

- We roll a die until we observed all possible results (from 1 to 6). What is the expected number of rolls?
- Each box of cereals contains a coupon, and there are n different types of coupons. How many boxes need to be bought on expectation to collect all n

coupons?

- X_1 — number of trails until collecting the first coupon ($X_1 = 1$)
- X_2 — number of additional trials (after collecting the first coupon) until collecting the second coupon (different than the first one)
- X_i — number of additional trials until collecting the ith coupon.
- Y — number of trails until collecting all n coupons

$$Y = X_1 + X_2 + \ldots + X_n$$

- The chance of getting a new coupon in a single trial is the number of possible new coupons divided by the number of all coupons, or $\frac{n-(i-1)}{n}$.
- We draw until we get the first new coupon (after getting the (i-1)th coupon), so $X_i \sim G_1(\frac{n-i+1}{n})$.
- Then $EX_i = \frac{1}{p_i} = \frac{n}{n-i+1}$
- $EY = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \ldots + \frac{n}{2} + \frac{n}{1}$
- $EY = n \cdot H_n$ where H_n is the harmonic number (partial sum of the harmonic series)
- $EY \simeq n \ln n$

Covariance

- The quantity:

$$C(X, Y) = E((X - EX)(Y - EY))$$

- is called the covariance of random variables X and Y.
- Covariance measures the relation between two random variables. A positive covariance means that both variables tend to be high or low at the same time. A negative covariance means that when one variable is high, the other tends to be low.

- $C(X, X) = D^2(X)$
- $C(X, -X) = -D^2(X)$
- If $C(X, Y) = 0$ then X and Y are called uncorrelated.
- $C(X, Y) = E(XY) - (EX)(EY)$

Variance of a sum and difference of random variables

$$D^2(X \pm Y) = D^2(X) \pm 2C(X, Y) + D^2(Y)$$

Cauchy-Schwarz inequality

- If $Y = cX$ for some constant $c \in \mathbb{R}$ then:

$$(E(XY))^2 \leq E(X^2)E(Y^2)$$

- It implies that:

$$|C(X, Y)| \leq D(X)D(Y)$$

- The above equality holds true iff one of the variables is a linear function of the other, e.g. $Y = aX + b$

Correlation coefficient

- A quantity:

$$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)}$$

- is called the correlation coefficient between random variables X and Y.
- $\rho(X, Y) \in [-1, 1]$
- $\rho(X, Y) \in \{-1, 1\}$ iff one variable is a linear function of the other.

- The correlation coefficient is the normalized covariance.

Independent random variables

- Random variables X and Y are called independent if:

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

- If $P(Y \in B) > 0$ then $P(X \in A|Y \in B) = P(X \in A)$
- Discrete random variables X and Y are independent if and only if:

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

- For all x, y.
- Random variables $X_1, \ldots, X_n$ are independent if:

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdot \ldots \cdot P(X_n \in A_n)$$

- For discrete random variables:

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot \ldots \cdot P(X_n = x_n)$$

- If X and Y are independent, then so are $g(X)$ and $h(Y)$ for any functions g and h. The same holds for any number of variables and for functions of multiple random variables.

Expectation of the product of independent random variables

- If random variables X and Y are independent then it holds:

$$E(XY) = (EX)(EY)$$

- If random variables $X_1, X_2, \ldots, X_n$ are independent then it holds:

$$E(X_1 \cdot X_2 \cdot \ldots \cdot X_n) = (EX_1) \cdot (EX_2) \cdot \ldots \cdot (EX_n)$$

Covariance of independent random variables

- If random variables X and Y are independent then it holds:

$$C(X, Y) = 0$$

- Therefore it also holds that:

$$D^2(X \pm Y) = D^2(X) + D^2(Y)$$

Summary

- For any random variables:

$$\begin{aligned} E(X+Y) &= EX + EY \\ D^2(X \pm Y) &= D^2(X) \pm 2C(X,Y) + D^2(Y) \end{aligned}$$

- For independent random variables:

$$\begin{aligned} E(XY) &= (EX)(EY) \\ C(X,Y) &= 0 \\ D^2(X \pm Y) &= D^2(X) + D^2(Y) \end{aligned}$$

Distribution of a sum of independent random variables

- Assume X and Y are independent. What is the distribution of $Z = X + Y$?

$$P(Z=z)=\sum_{x,y:\ x+y=z} P(X=x)P(Y=y)$$

Binomial distribution

- Consider a Bernoulli scheme with n trials with the success probability p.
- $X_i \in \{0,1\}$ — the outcome in the ith trial
- Variables X_i have Bernoulli distribution $B(p)$ and are independent
- $Y = X_1 + X_2 + \ldots + X_n$ — number of successes in n trials
- Y has a binomial distribution $B(n,p)$

$$EX_i = p$$
$$D^2(X_i) = p(1-p)$$

$$EY = EX_1 + EX_2 \ldots + EX_n = np$$
$$D^2(Y) = D^2(X_1) + \ldots + D^2(X_n) = np(1-p)$$