Seminar Nr. 6, Numerical Characteristics of Random Variables

Theory Review

Expectation:

- if $X \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}$ is discrete, then $E(X) = \sum\limits_{i \in I} x_i p_i$.

- if X is continuous with pdf f, then $E(X) = \int\limits_{\mathbb{R}} x f(x) dx$.

Variance: $V(X) = E\left((X - E(X))^2\right) = E\left(X^2\right) - (E(X))^2$.

Standard Deviation: $\sigma(X) = \operatorname{Std}(X) = \sqrt{V(X)}$.

Moments:

- **moment of order k**: $\nu_k = E\left(X^k\right)$.
- **absolute moment of order k**: $\underline{\nu_k} = E\left(|X|^k\right)$.
- **central moment of order k**: $\mu_k = E\left((X - E(X))^k\right)$.

Properties:

1. $E(aX + b) = aE(X) + b, V(aX + b) = a^2 V(X)$
2. $E(X + Y) = E(X) + E(Y)$
3. if X and Y are independent, then $E(XY) = E(X)E(Y)$ and $V(X + Y) = V(X) + V(Y)$
4. if $h : \mathbb{R} \to \mathbb{R}$ is a measurable function, X a random variable;

- if X is discrete, then $E(h(X)) = \sum\limits_{i \in I} h(x_i) p_i$
- if X is continuous, then $E(h(X)) = \int\limits_{\mathbb{R}} h(x) f(x) dx$

Covariance: $\operatorname{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$

Correlation Coefficient: $\rho(X, Y) = \dfrac{\operatorname{cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$

Properties:

1. $\operatorname{cov}(X, Y) = E(XY) - E(X)E(Y)$
2. $V\left(\sum\limits_{i=1}^{n} a_i X_i\right) = \sum\limits_{i=1}^{n} a_i^2 V(X_i) + 2 \sum\limits_{1 \le i < j \le n} a_i a_j \operatorname{cov}(X_i, X_j)$
3. X, Y independent => $\operatorname{cov}(X, Y) = \rho(X, Y) = 0$ (X and Y are *uncorrelated*)
4. $-1 \le \rho(X, Y) \le 1$; $\rho(X, Y) = \pm 1 \iff \exists\, a, b \in \mathbb{R},\ a \neq 0$ s.t. $Y = aX + b$

Let (X, Y) be a continuous random vector with pdf $f(x, y)$, let $h : \mathbb{R}^2 \to \mathbb{R}^2$ a measurable function, then

$$E(h(X, Y)) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} h(x, y) f(x, y) dx dy$$

---

1. Every day, the number of network blackouts has the following pdf

$$X \begin{pmatrix} 0 & 1 & 2 \\ 0.7 & 0.2 & 0.1 \end{pmatrix}.$$

A small internet trading company estimates that each network blackout costs them \$500.

a) How much money can the company expect to lose each day because of network blackouts?

b) What is the standard deviation of the company's daily loss due to blackouts?

Solution:

Variable X represents the number of daily network blackouts and each network blackout costs \$500. Then the company's daily loss due to blackouts is the variable

$$Y = 500X.$$

a) We want the *expected* daily loss due to blackouts, so $E(Y)$. First, let us compute $E(X)$. Since X is a *discrete* random variable, we compute its expectation as a sum.

$$E(X) \;=\; \sum_{i\in I} x_i p_i \;=\; 0\cdot 0.7 + 1\cdot 0.2 + 2\cdot 0.1 \;=\; 0.4.$$

Then, by the properties of expectation,

$$E(Y) \;=\; E(500X) \;=\; 500E(X) \;=\; 500\cdot 0.4 \;=\; 200 \text{ dollars.}$$

b) Now we want

$$\sigma(Y) \;=\; \sqrt{V(Y)}.$$

Again, let us first compute $V(X)$. For that, we need $E(X^2)$. The pdf of X^2 is

$$X^2 \begin{pmatrix} 0 & 1 & 4 \\ 0.7 & 0.2 & 0.1 \end{pmatrix},$$

so

$$\begin{aligned} E(X^2) &= 0.2 + 0.4 \;=\; 0.6, \\ V(X) &= E(X^2) - (E(X))^2 \;=\; 0.6 - 0.16 \;=\; 0.44. \end{aligned}$$

Then, by the properties of variance,

$$\begin{aligned} V(Y) &= V(500X) \;=\; (500)^2 V(X) \;=\; 250000\cdot 0.44 \;=\; 110000 \text{ dollars}^2 \text{ and} \\ \sigma(Y) &= \sqrt{V(Y)} \;=\; 331.6625 \text{ dollars.} \end{aligned}$$

2. (Refer to Problem 5 in Sem. 3) About ten percent of computer users in a public library do not close Windows properly. On the average, how many users *do* close Windows properly before someone *does not*?

Solution:

Let X denote the number of users that close Windows properly before someone does not. If "success" means "a person *does not* close Windows properly", then X has a Geometric distribution with parameter $p = 0.1$. We want the "average" of that number, i.e. the mean value $E(X)$. For the *Geo*(p) distribution, the mean value is $\frac{q}{p}$. Thus,

$$E(X) \;=\; \frac{q}{p} \;=\; \frac{0.9}{0.1} \;=\; 9.$$

3. (Refer to Problem 1 in Sem. 5) The lifetime, in years, of some electronic component is a random variable with density

$$f(x) = \begin{cases} \dfrac{3}{x^4}, & \text{for} \quad x \ge 1 \\ 0, & \text{for} \quad x < 1. \end{cases}$$

How many years, on the average, can we expect that electronic equipment to last?

Solution:

The variable X is the lifetime of the component, we want to know how long can we *expect* it to last, so, again, $E(X)$. This is a *continuous* random variable, so we compute its expectation as an integral. We have

$$\begin{aligned} E(X) &= \int\limits_{\mathbb{R}} x f(x)\,dx \;=\; 3\int\limits_{1}^{\infty} x \cdot \frac{1}{x^4}\,dx \\ &= 3\int\limits_{1}^{\infty} x^{-3}\,dx \;=\; -\frac{3}{2}\frac{1}{x^2}\Big|_1^{\infty} \;=\; \frac{3}{2}, \end{aligned}$$

so about a year and a half.

4. (Optimal portfolio) A businessman wants to invest \$600 and has two companies to choose from, company A, where shares cost \$20 each and company B, where shares cost \$30 per share. The market analysis shows that for company A the return per share is distributed as follows: lose \$1 with probability 0.2, win \$2 with probability 0.6, or win/lose nothing. For company B: lose \$1 with probability 0.3, win \$3 with probability 0.6, or win/lose nothing. The returns from the two companies are independent. In order to maximize the expected return and minimize the risk, which way is better to invest:
a) all money in company A;
b) all money in company B;
c) half the amount in each?

Solution:

So, we want to maximize the *expected* return, that means the expected value of the return and minimize the risk. How do we quantify the "risk"? Once we know the average (expected) value of the return, "risk" would be the amount of *variability* from that expected return, i.e. its variance (or standard deviation). For an "optimal" portfolio, we will want *high* expected return and *low* variance.
Let A denote the actual (random) return from each share of company A and B the same for company B. Then their pdf's are

$$A\begin{pmatrix} -1 & 0 & 2 \\ 0.2 & 0.2 & 0.6 \end{pmatrix} \text{ and } B\begin{pmatrix} -1 & 0 & 3 \\ 0.3 & 0.1 & 0.6 \end{pmatrix}.$$

The pdf's of A^2 and B^2 (which will be needed for the computation of the variances) are

$$A^2\begin{pmatrix} 0 & 1 & 4 \\ 0.2 & 0.2 & 0.6 \end{pmatrix} \text{ and } B^2\begin{pmatrix} 0 & 1 & 9 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}.$$

We have

$$\begin{aligned} E(A) &= -0.2+1.2 = 1, & E(B) &= -0.3+1.8 = 1.5, \\ E(A^2) &= 0.2+2.4 = 2.6, & E(B^2) &= 0.3+5.4 = 5.7, \\ V(A) &= 2.6-1^2 = 1.6, & V(B) &= = 5.7-(1.5)^2 = 3.45. \end{aligned}$$

Now we can make a comparison between the three investments, by looking at their expected values and their variances.
a) Investing **all money in company A**, i.e. \$600 at \$20 per share, means buying 30 shares from company A. So the return (profit) from this investment is the random variable $30A$, for which
$E(30A) = 30E(A) = 30$ (the expected return)
$V(30A) = 900V(A) = 1440$ (the variance of the return, i.e. the risk of the investment)

b) Investing **all money in company B**, i.e. \$600 at \$30 per share, means buying 20 shares from company B. The return from this investment is the random variable $20B$, for which
$E(20B) = 20E(B) = 30$ (the expected return)
$V(20B) = 400V(B) = 1380$ (the variance of the return, i.e. the risk of the investment)

c) Finally, investing **half the amount in each**, i.e. \$300 at \$20 per share, means buying 15 shares of company A stock and \$300 at \$30 per share, means buying 10 shares of company B stock. In this case, the return from this investment is the random variable $15A + 10B$, for which
$E(15A + 10B) = 15E(A) + 10E(B) = 30$ (the expected return)
and, because A and B are independent,
$V(15A + 10B) = 225V(A) + 100V(B) = 705$ (the variance of the return, i.e. the risk of the investment)

So **all** three proposed portfolios have the same expected return (which should be no surprise, since each share of each company is expected to return $1/20$ or $1.5/30$, which is 5%), but the **third** one, with the lowest variance is the **least risky**.
This is why financiers and brokers always say "in order to minimize the risk, D I V E R S I F Y the portfolio"..

5. (Reduced Variables). Let X be a random variable with mean $E(X)$ and standard deviation $\sigma(X) = \sqrt{V(X)}$. Find the mean and variance of $Y = \dfrac{X - E(X)}{\sigma(X)}$.

Solution:
This is a simple linear transformation, that can be applied to every random variable (discrete or continuous) that has a mean value and a variance (i.e. they are finite):

$$Y = \frac{1}{\sigma(X)}X - \frac{E(X)}{\sigma(X)} = aX + b.$$

The "reduction" refers to the mean value and variance of the reduced variable. Recall the properties

$$\begin{aligned} E(aX + b) &= aE(X) + b, \\ V(aX + b) &= a^2V(X). \end{aligned}$$

So

$$\begin{aligned} E(Y) &= \frac{1}{\sigma(X)}E(X) - \frac{E(X)}{\sigma(X)} = 0, \\ V(Y) &= \frac{1}{\sigma^2(X)}V(X) = 1. \end{aligned}$$

Y is called the **reduced variable of** X.

Recall that for a Normal $N(\mu, \sigma)$ variable, $E(X) = \mu, V(X) = \sigma^2$.
So, returning to reduced variables, in particular, if $X \in N(\mu, \sigma)$, then $\dfrac{X - \mu}{\sigma} \in N(0, 1)$ (Standard or Reduced Normal). We see many such reduced variables in Statistics:

$$\frac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}, \text{ or, more generally, } \frac{\overline{\theta} - \theta}{\sigma_{\overline{\theta}}}.$$

6. The joint density function of the vector (X, Y) is $f(x, y) = x + y,\ (x, y) \in [0, 1] \times [0, 1]$. Find
a) the means and variances of X and Y;
b) the correlation coefficient $\rho(X, Y)$.

Solution:
To compute *all* of these numerical characteristics, we use the formula

$$E\left(h(X, Y)\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f(x, y)dxdy.$$

a)

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy \\
&= \int_{0}^{1}\int_{0}^{1}(x^2+xy)dydx \;=\; \int_{0}^{1}\left(x^2\cdot y + x\cdot\frac{1}{2}y^2\right)\Bigg|_{y=0}^{y=1} dx \\
&= \int_{0}^{1}\left(x^2+\frac{1}{2}x\right)dx \;=\; \left(\frac{1}{3}x^3+\frac{1}{4}x^2\right)\Bigg|_{0}^{1} \;=\; \frac{7}{12},
\end{aligned}$$

and by symmetry, $E(Y)=\frac{7}{12}$, also.

$$E\left(X^2\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2f(x,y)dxdy = \int_{0}^{1}\int_{0}^{1}(x^3+x^2y)dxdy = \frac{5}{12} = E\left(Y^2\right),$$

again, by symmetry.
That means the variances are also equal.

$$V(X) = V(Y) = \frac{5}{12}-\left(\frac{7}{12}\right)^2 = \frac{11}{12^2}.$$

b) This formula is *especially* useful for computing the covariance, since it avoids the (many times) complicated calculation of the pdf of the variable XY. We have

$$\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy \\
&= \int_{0}^{1}\int_{0}^{1}(x^2y+xy^2)dxdy \;=\; \frac{1}{3},
\end{aligned}$$

so

$$\begin{aligned}
\text{cov}(X,Y) &= E(XY)-E(X)E(Y) \\
&= \frac{1}{3}-\left(\frac{7}{12}\right)^2 \;=\; -\frac{1}{12^2} \\
\rho(X,Y) &= \frac{\text{cov}(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} \\
&= -\frac{\frac{1}{12^2}}{\frac{11}{12^2}} \;=\; -\frac{1}{11}.
\end{aligned}$$

The value of $|\rho(X,Y)|$ is so close to 0, which means there is very weak (almost nonexistent) linear relationship between X and Y.

7. Let X be a discrete random variable with pdf $X\begin{pmatrix} -1 & 0 & 1 \\ \sin^2 a & \cos 2a & \sin^2 a \end{pmatrix}$, $a \in \left(0, \frac{\pi}{4}\right)$. For any $k \in \mathbb{N}^*$, let $Y_k = X^{2k-1}$ and $Z_k = X^{2k}$. Find $\rho(Y_k, Z_k)$. (In particular, X and X^2 are uncorrelated, but *not* independent).

Solution:

Recall that

$$X, Y \text{ are independent } => X, Y \text{ are uncorrelated, i.e. } \rho(X,Y) = 0.$$

That makes sense, since independence means there is no relationship of *any kind* between the variables, including linear. Obviously, independence is a much stronger condition than uncorrelation. A linear relationship does not exist, but some *other* form of relationship may exist. This exercise is an example in that sense, variables that are uncorrelated, but *not* independent.

Notice that for any $k \in \mathbb{N}^*$, $Y_k = X^{2k-1}$ has the same pdf as

$$X\begin{pmatrix} -1 & 0 & 1 \\ \sin^2 a & \cos 2a & \sin^2 a \end{pmatrix}$$

and $Z_k = X^{2k}$ has the same pdf as

$$X^2\begin{pmatrix} 0 & 1 \\ \cos 2a & 2\sin^2 a \end{pmatrix}.$$

Then

$$\begin{aligned} E\left(X^{2k-1}\right) &= -\sin^2 a + \sin^2 a = 0, \\ E\left(X^{2k}\right) &= 2\sin^2 a, \\ E\left(X^{2k-1}X^{2k}\right) &= E\left(X^{4k-1}\right) = 0, \end{aligned}$$

so

$$\operatorname{cov}(Y,Z) = \rho(Y,Z) = 0.$$

In particular, for $k = 1$, we have that $\rho\left(X, X^2\right) = 0$, so they are uncorrelated, but obviously, not independent.