Seminar 5

Recap of random variables

Consider probabiliy space $(\Omega, \mathcal{F}, \mathbb{P})$. A **random variable** is a measurable function $X : \Omega \to \mathbb{R}$ from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Measurable function $X : \Omega \to \mathbb{R}$ from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. It means that the pre-image of any set A in $\mathcal{B}(\mathbb{R})$ belongs to $\mathcal{F}$:

$$\forall A \in \mathcal{B}(\mathbb{R}) \Longrightarrow X^{-1}(A) \in \mathcal{F}$$

Recap of distributions

We will call the image μ of measure $\mathbb{P}$ through the mapping X **the distribution** of r.v. X and write $X \sim \mu$:

$$\mu(A) = \mathbb{P}(X^{-1}(A))$$

Any probabilistic measure (hence any probability distribution) can be decomposed into sum of three types of measures:

- Discrete
- Singular
- Absolutely continuous

Normally, the distributions fall into just one category and you never encounter singular distributions.

Functions describing distributions

- For any distribution we have **cumulative distribution function** (CDF) $F_X(x) = \mathbb{P}(X < x)$
- For discrete distributions we have **probability mass function** (PMF) $\mathbb{P}_X(x) = \mathbb{P}(X = x)$
- For continuous distributions we have **probability density function** (PDF) $f_X(x) = F'_X(x)$

Functions of random variables

Random variables transform like functions, i.e. if $Y = \varphi(X)$, then $Y(\omega) = \varphi(X(\omega))$.

For a smooth φ, the density will be:

$$f_Y(y) = \sum_{\varphi(x)=y} \frac{f_X(x)}{|\varphi'(x)|}$$

Mathematical expectation

Mathematical expectation generalizes the concept of mean. Consider probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variable $X : \Omega \to \mathbb{R}$. Then expected value of X is

$$\mathbb{E}\left[X\right] = \int_\Omega X(\omega) d\mathbb{P}(\omega) = \int_\mathbb{R} x d\mu(x)$$

- If X is discrete, then

$$\mathbb{E}\left[X\right] = \sum_k x_k \mathbb{P}(X = x_k)$$

- If X is continuous, then

$$\mathbb{E}\left[X\right] = \int_{-\infty}^{+\infty} x f_X(x) dx$$

It may be the case that $\mathbb{E}\left[X\right] = \pm\infty$ or even does not exist.

Example 1

We roll a die and r.v. X is the score of a roll. What is $\mathbb{E}\left[X\right]$?

Solution 1

$$\mathbb{E}\left[X\right] = \sum_{k=1}^{6} k \cdot \mathbb{P}(X = k) = \frac{1}{6} \sum_{k=1}^{6} k = \frac{7}{2}$$

Example 2

We flip a non-symmetric coin and X is the r.v. for heads, $X \sim Be(p)$. What is $\mathbb{E}\left[X\right]$?

Solution 2

$$\mathbb{E}\left[X\right] = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) = p$$

Example 3

Consider discrete r.v. X with distribution $\mathbb{P}(X = 2^n) = 2^{-n}$. What is $\mathbb{E}\left[X\right]$?

Solution 3

$$\mathbb{E}\left[X\right] = \sum_n 2^n 2^{-n} = \infty$$

Example 4

Consider discrete r.v. X with distribution $\mathbb{P}(X = 2^n) = \mathbb{P}(X = -2^n) = 2^{-n-1}$. What is $\mathbb{E}\left[X\right]$?

Solution 4

Expectation of r.v. X exists if and only if $\mathbb{E}\left[|X|\right] < \infty$

Example 5

Consider X with **Poisson distribution** $X \sim Pois(\lambda)$:

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

What is $\mathbb{E}\left[X\right]$?

Solution 5

$$\mathbb{E}\left[X\right] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k-1)!} =$$
$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k+1}}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Properties of expectation

Consider r.v.s X and Y with finite expectations. Then,

1. For any constants a and b it holds $\mathbb{E}\left[aX+b\right] = a\mathbb{E}\left[X\right] + b$
2. $\mathbb{E}\left[X+Y\right] = \mathbb{E}\left[X\right] + \mathbb{E}\left[Y\right]$
3. If $X \leqslant Y$ a.s., then $\mathbb{E}\left[X\right] \leqslant \mathbb{E}\left[Y\right]$
4. If $X \perp Y$, then $\mathbb{E}\left[XY\right] = \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$

Example 6

Consider X with binomial distribution $X \sim Bi(n,p)$. What is $\mathbb{E}\left[X\right]$?

Solution 6

- We know that $X = \sum_{k=1}^n X_k$, where $X_k \sim Be(p)$
- We know that $\mathbb{E}\left[X_k\right] = p$
- Then, $\mathbb{E}\left[X\right] = \sum_{k=1}^n \mathbb{E}\left[X_k\right] = np$

Expectation of a function of a random variable

Consider $Y = \varphi(X)$, then its expectation is

$$\mathbb{E}\left[Y\right] = \int_\Omega \varphi(X(\omega))d\mathbb{P}_X(\omega) = \int_{-\infty}^{\infty} \varphi(x)d\mu(x)$$

If additionally the following integral exists

$$\int_{-\infty}^{\infty} |\varphi(x)|dF_X(x) < \infty$$

Then,

$$\mathbb{E}\left[Y\right] = \int_{-\infty}^{\infty} \varphi(x)f_X(x)dx$$

Variance

We call **variance** the following quantity of a r.v. X with finite expectation:

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

Example 7

We flip a non-symmetric coin and X is the r.v. for heads, $X \sim Be(p)$. What is $\mathbb{V}\mathrm{ar}\left(X\right)$?

Solution 7

1. We know the formula

$$\mathbb{V}\mathrm{ar}\left(X\right) = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)^2\right]$$

2. We know $\mathbb{E}\left[X\right]$

$$\mathbb{V}\mathrm{ar}\left(X\right) = \mathbb{E}\left[\left(X - p\right)^2\right] = \mathbb{E}\left[X^2 - 2pX + p^2\right]$$

3. We know that expectation is linear

$$\mathbb{V}\mathrm{ar}\left(X\right) = \mathbb{E}\left[X^2\right] - 2p\mathbb{E}\left[X\right] + p^2 = \mathbb{E}\left[X^2\right] - p^2$$

4. For $Y = X^2$ we can compute

$$\mathbb{E}\left[Y\right] = 0 \cdot \mathbb{P}(Y = 0) + 1 \cdot \mathbb{P}(Y = 1) = \mathbb{P}(Y = 1) = \mathbb{P}(X^2 = 1) = \mathbb{P}(X = 1) = p$$

5. Finally,

$$\mathbb{V}\mathrm{ar}\left(X\right) = p - p^2 = p(1 - p)$$

Properties of variance

1. $\mathbb{V}\mathrm{ar}\left(X\right) \geqslant 0$ and $\mathbb{V}\mathrm{ar}\left(X\right) = 0$ if and only if $X = const$ a.s.
2. If holds

$$\mathbb{V}\mathrm{ar}\left(X\right) = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2$$

3. It holds

$$\mathbb{V}\mathrm{ar}\left(aX + b\right) = b^2\mathbb{V}\mathrm{ar}\left(X\right)$$

4. If $X \perp Y$, it holds

$$\mathbb{V}\mathrm{ar}\left(X + Y\right) = \mathbb{V}\mathrm{ar}\left(X\right) + \mathbb{V}\mathrm{ar}\left(Y\right)$$

Example 8

Consider X with binomial distribution $X \sim Bi(n, p)$. What is $\mathbb{V}\mathrm{ar}\left(X\right)$?

Solution 8

- We know that $X = \sum_{k=1}^{n} X_k$, where $X_k \sim Be(p)$
- We know that $\mathbb{V}\mathrm{ar}\left(X_k\right) = p(1-p)$
- Then, $\mathbb{V}\mathrm{ar}\left(X\right) = \mathbb{V}\mathrm{ar}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}\mathrm{ar}\left(X_k\right) = np(1-p)$

Moments of distribution

$\mathbb{E}\left[X^k\right]$ is called k-th moment of r.v. X.

We say that k-th moment is finite if $\mathbb{E}\left[X^k\right] < \infty$.

If k-th moment is finite, then all moments $m < k$ are finite as well.

Jensen's inequality

Consider r.v. X with $\mathbb{E}[X] < \infty$ and convex function $g(\cdot)$, then

$$\mathbb{E}\left[g(X)\right] \geqslant g\left(\mathbb{E}\left[X\right]\right)$$

Example 9

Prove Jensen's inequality for special case of $g(x) = x^2$

Cauchy-Schwarz inequality

Consider r.v. X with $\mathbb{E}\left[X^2\right] < \infty$, then

$$\left|\mathbb{E}\left[XY\right]\right| \leqslant \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}$$

Covariance

Covariance of two random variables X and Y is defined as

$$\mathrm{cov}(X, Y) = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)\left(Y - \mathbb{E}\left[Y\right]\right)\right] = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$$

From Cauchy-Schwarz inequality,

$$\mathrm{cov}(X, Y) \leqslant \sqrt{\mathbb{V}\mathrm{ar}\left(X\right)\mathbb{V}\mathrm{ar}\left(Y\right)}$$

Correlation

If $X \perp Y$, $\operatorname{cov}(X, Y) = 0$. The converse is not true. Regardless, covariance is often used to measure the dependency between random variables. It is not handy to use, so instead a **correlation coefficient** is proposed:

$$r_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{\mathbb{V}\mathrm{ar}\,(X)\,\mathbb{V}\mathrm{ar}\,(Y)}}$$

Note that $-1 \leqslant r_{XY} \leqslant 1$.