Review of Selected Topics in Probability Expectation

Christoforos Raptopoulos

Lecture 5

Expectation

...or mean value, expected value, 1st moment.

Definition (Expectation)

The expectation of a random variable X is the number

$$\mathbb{E}[X] \stackrel{def}{=} \sum_{x \in \mathcal{A}_X} x \Pr(X = x) \qquad \text{(if } X \text{ is discrete)} \qquad (1)$$

$$\stackrel{def}{=} \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(if } X \text{ is continuous)} \qquad (2)$$

Expectation

...or mean value, expected value, 1st moment.

Definition (Expectation)

The expectation of a random variable X is the number

$$\mathbb{E}[X] \stackrel{def}{=} \sum_{x \in \mathcal{A}_X} x \Pr(X = x) \qquad \text{(if } X \text{ is discrete)} \tag{1}$$

$$\stackrel{def}{=} \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(if } X \text{ is continuous)} \tag{2}$$

Quiz: $\mathbb{E}[X] \in \mathcal{A}_X$? Answer:

Expectation

...or mean value, expected value, 1st moment.

Definition (Expectation)

The expectation of a random variable X is the number

$$\mathbb{E}[X] \stackrel{def}{=} \sum_{x \in \mathcal{A}_X} x \Pr(X = x) \qquad \text{(if } X \text{ is discrete)} \qquad (1)$$

$$\stackrel{def}{=} \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(if } X \text{ is continuous)} \qquad (2)$$

Quiz: $\mathbb{E}[X] \in \mathcal{A}_X$? Answer: No, it is just a *weighted sum* of the values of X!

Definitions and Properties of Expectation

- $\mathbb{E}[c \cdot X] = c\mathbb{E}[X]$, for any constant c.

- Theorem (Linearity of Expectation)
  For any random variables X, Y, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

  - For any real function $g(\cdot)$ defined on $\mathcal{A}_X$,
    $\mathbb{E}[g(X)] = \sum_{x \in \mathcal{A}_X} g(x) \Pr(X = x)$.

  - For any two independent random variables X, Y,
    $\mathbb{E}[X \cdot Y] = \mathbb{E}[X]\mathbb{E}[Y]$.

  - There is at least one value $a_l \in \mathcal{A}_X$ such that $a_l \geq \mathbb{E}[X]$.

First Order Concentration - Markov's Inequality

Question: Suppose we flip a coin 50 times and let X be the number of times it turned up heads. How much is the probability that $X > 47$?

First Order Concentration - Markov's Inequality

Question: Suppose we flip a coin 50 times and let X be the number of times it turned up heads. How much is the probability that $X > 47$?

Theorem (Markov's Inequality)

For any positive random variable X, and $t > 0$

$$\Pr(X \geq t) \leq \frac{\mathbb{E}[X]}{t}. \quad (3)$$

Proof (discrete case).

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in \mathcal{A}_X} x \Pr(X = x) \geq \sum_{x \geq t} x \Pr(X = x) \\ &\geq \sum_{x \geq t} t \Pr(X = x) = \end{aligned}$$

First Order Concentration - Markov's Inequality

Question: Suppose we flip a coin 50 times and let X be the number of times it turned up heads. How much is the probability that $X > 47$?

Theorem (Markov's Inequality)

For any positive random variable X, and $t > 0$

$$\Pr(X \geq t) \leq \frac{\mathbb{E}[X]}{t}. \tag{3}$$

Proof (discrete case).

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in \mathcal{A}_X} x \Pr(X = x) \geq \sum_{x \geq t} x \Pr(X = x) \\ &\geq \sum_{x \geq t} t \Pr(X = x) = t \Pr(X \geq t). \quad \square \end{aligned}$$

Variance

Example: Let $W = 0$ with probability 1,

$$Y = \begin{cases} +1 & \text{,with probability } \frac{1}{2} \\ -1 & \text{,with probability } \frac{1}{2}. \end{cases}$$

and

$$X = \begin{cases} +100 & \text{,with probability } \frac{1}{2} \\ -100 & \text{,with probability } \frac{1}{2}. \end{cases}$$

These are different, but have the same expectation.

Variance

Example: Let $W = 0$ with probability 1,

$$Y = \begin{cases} +1 & \text{,with probability } \frac{1}{2} \\ -1 & \text{,with probability } \frac{1}{2}. \end{cases}$$

and

$$X = \begin{cases} +100 & \text{,with probability } \frac{1}{2} \\ -100 & \text{,with probability } \frac{1}{2}. \end{cases}$$

These are different, but have the same expectation.

Definition (Variance)

The variance of a random variable X is the number

$$Var(X) \stackrel{def}{=} \mathbb{E}[(X - \mu)^2] \tag{4}$$

where $\mu = \mathbb{E}[X]$. The typical deviation of X is defined as

$$\sigma \stackrel{def}{=} \sqrt{Var(X)}. \tag{5}$$

Properties of Variance

- $Var(X) \geq 0$.
- $Var(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$.
- $Var(aX + b) = a^2 Var(X)$, for any constants a, b.
- If X, Y are independent, then $Var(X + Y) = Var(X) + Var(Y)$. Note: The inverse does not hold!

Covariance

Definition (Covariance)

The covariance of two random variables X, Y is

$$\begin{aligned} Cov(X,Y) &\stackrel{def}{=} \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \end{aligned} \tag{6}$$

,where $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$.

Intuition: The covariance is a measure of the dependency of two variables. It can take negative values (can you find an example?).

Properties of Covariance

- If X is independent of Y, $Cov(X, Y) = 0$. Note: The inverse does not hold! (Homework: Find an example!)
- $Cov(X, Y) = Cov(Y, X)$ (symmetry).
- $Cov(X, X) = Var(X)$.
- (variance of sum)

$$
\begin{aligned}
Var\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i,j} Cov(X_i, X_j) \\
&= \sum_{i} Var(X_i) + \sum_{i} \sum_{j \neq i} Cov(X_i, X_j).
\end{aligned}
$$

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

For any random variable X and any $k > 0$

$$\Pr(|X - \mathbb{E}[X]| \geq k) \leq \frac{Var(X)}{k^2}. \tag{7}$$

Proof. Similar to the proof of Markov's inequality. □

Note: Chebyshev's inequality gives stronger bounds than Markov's inequality, at the cost of more complex computations. Intuitively, the more information we have from higher moments (i.e. $\mathbb{E}[X^i]$, for $i \in \mathbb{N}$), the more we can say for a distribution (more on this later...).

Quiz

For some fixed $k > 0$, let

$$X = \begin{cases} k & \text{,with probability } \frac{1}{2k^2} \\ 0 & \text{,with probability } 1 - \frac{1}{k^2} \\ -k & \text{,with probability } \phi. \end{cases}$$

What are is the exact value and the upper bounds we can get for $\Pr(|X| \geq k)$ using Markov's and Chebyshev's inequality respectively?

(a) $\frac{1}{2}, 1$ and $\frac{1}{k^2}$

(b) $\frac{1}{k^2}$, inconclusive and $\frac{1}{k^2}$

(c) $\frac{1}{2}, 0$ and $\frac{1}{2}$

(d) none of the above

Quiz

For some fixed $k > 0$, let

$$X = \begin{cases} k & \text{,with probability } \frac{1}{2k^2} \\ 0 & \text{,with probability } 1 - \frac{1}{k^2} \\ -k & \text{,with probability } \phi. \end{cases}$$

What are is the exact value and the upper bounds we can get for $\Pr(|X| \geq k)$ using Markov's and Chebyshev's inequality respectively?

(a) $\frac{1}{2}, 1$ and $\frac{1}{k^2}$

(b) $\frac{1}{k^2}$, inconclusive and $\frac{1}{k^2}$

(c) $\frac{1}{2}, 0$ and $\frac{1}{2}$

(d) none of the above

Answer: (b)

Conditional Expectation

Definition (Conditional Expectation)

Let X, Y two random variables. The conditional expectation of X, given $Y = y$ is the function of y

$$\mathbb{E}[X|Y=y] \stackrel{def}{=} \sum_{x \in \mathcal{A}_X} x \cdot \Pr(X=x|Y=y) \qquad \text{(if } X \text{ is discrete)}$$

$$\stackrel{def}{=} \int_{-\infty}^{\infty} x \cdot f(x|y)dx \qquad \text{(if } X \text{ is continuous)}$$

Conditional Expectation

Definition (Conditional Expectation)

Let X, Y two random variables. The conditional expectation of X, given $Y = y$ is the function of y

$$\mathbb{E}[X|Y = y] \stackrel{def}{=} \sum_{x \in \mathcal{A}_X} x \cdot \Pr(X = x|Y = y) \qquad \text{(if } X \text{ is discrete)}$$

$$\stackrel{def}{=} \int_{-\infty}^{\infty} x \cdot f(x|y) dx \qquad \text{(if } X \text{ is continuous)}$$

- If X, Y are independent, then $\mathbb{E}[X|Y = y] = \mathbb{E}[X]$
- $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$ (Exercise: Prove this!)

Homework

Investigate the discrete distributions

- Binomial
- Geometric
- Poisson

and the continuous distributions

- Uniform
- Exponential

in terms of Expectation, Variance.

Further reading

S. Ross. A first course in probability:
Chapter 4, "Random Variables"
Chapter 5, "Continuous Random Variables"
Chapter 6, "Jointly Distributed Random Variables"
Chapter 7, "Properties of Expectation"