Lecture 16: Expected value, variance, independence and Chebyshev inequality

Expected value, variance, and Chebyshev inequality. If X is a random variable recall that the *expected value of* X, $E[X]$ is the average value of X

$$\textbf{Expected value of } X : \quad E[X] = \sum_{\alpha} \alpha \, P(X = \alpha)$$

The expected value measures only the average of X and two random variables with the same mean can have very different behavior. For example the random variable X with

$$P(X = +1) = 1/2 \,, \quad P(X = -1) = 1/2$$

and the random variable Y with

$$[P(X = +100) = 1/2 \,, \quad P(X = -100) \;=\; 1/2$$

have the same mean

$$E[X] = E[Y] \;=\; 0$$

To measure the "spread" of a random variable X, that is how likely it is to have value of X very far away from the mean we introduce the *variance of* X, denoted by $var(X)$. Let us consider the distance to the expected value i.e., $|X - E[X]|$. It is more convenient to look at the square of this distance $(X - E[X])^2$ to get rid of the absolute value and the variance is then given by

$$\textbf{Variance of } X \;:\; \mathrm{var}(X) \;=\; E\left[(X - E[X])^2\right]$$

We summarizesome elementary properties of expected value and variance in the following

Theorem 1. *We have*

1. *For any two random variables* X *and* Y, $E[X + Y] = E[X] + E[Y]$.
2. *For any real number* a, $E[aX] = aE[X]$.
3. *For any real number* a, $\mathrm{var}(aX) = a^2 \mathrm{var}(X)$.

Proof. For 1. one just needs to write down the definition. For 2. one notes that if X takes the value α with some probability then the random variable aX takes the value $a\alpha$ with the same probability. Finally for 3., we use 2. and we have

$$\mathrm{var}(X) \;=\; E\left[(aX - E[aX])^2\right] \;=\; E\left[a^2\,(X - E[X])^2\right] \;=\; a^2 E\left[(X - E[X])^2\right] \;=\; a^2\mathrm{var}(X)$$

∎

Example: The $0-1$ random variable. Suppose A is an event the random variable X_A is given by

$$X_A \;=\; \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

and let us write

$$p = P(A)$$

The we have

$$E[X_A] = 0 \times P(X_A = 0) + 1 \times P(X_A = 1) \;=\; 0 \times (1-p) + 1 \times p \;=\; p\,.$$

To compute the variance note that

$$X_A - E[X_A] \;=\; \begin{cases} 1-p & \text{if } A \text{ occurs} \\ -p & \text{otherwise} \end{cases}$$

and so

$$\mathrm{var}(X) \;=\; (-p)^2 \times P(X_A = 0) + (1-p)^2 \times P(X_A = 1) \;=\; p^2(1-p) + (1-p)^2 p \;=\; p(1-p)$$

In summary we have

The $0-1$ random variable

$$\begin{array}{ll} P(X=1) = p\,, & P(X=0) = (1-p) \\ E[X] = p\,, & \mathrm{var}(X) \;=\; p(1-p) \end{array}$$

∎

Chebyshev inequality: The Chebyshev inequality is a simple inequality which allows you to extract information about the values that X can take if you know only the mean and the variance of X.

Theorem 2. *We have*

1. **Markov inequality**. *If $X \geq 0$, i.e. X takes only nonnegative values, then for any $a > 0$ we have*
$$P(X \geq a) \;\leq\; \frac{E[X]}{\alpha}$$

2. **Chebyshev inequality**. *For any random variable X and any $\epsilon > 0$ we have*
$$P\left(|X - E[X]| \geq \epsilon\right) \leq \frac{\mathrm{var}(X)}{\epsilon^2}$$

Proof. Let us prove first Markov inequality. Pick a positive number a. Since X takes only nonnegative values all terms in the sum giving the expectations are nonnegative we have

$$E[X] \;=\; \sum_{\alpha} \alpha P(X = \alpha) \;\geq\; \sum_{\alpha \geq a} \alpha P(X = \alpha) \geq a \sum_{\alpha \geq a} P(X = \alpha) \;=\; a P(X \geq a)$$

and thus
$$P(X \geq a) \;\leq\; \frac{E[X]}{a}\,.$$
To prove Chebyshev we will use Markov inequality and apply it to the random variable

$$Y = (X - E[X])^2$$

which is nonnegative and with expected value

$$E[Y] \;=\; E\left[(X - E[X])^2\right] = \mathrm{var}(X)\,.$$

We have then

$$\begin{aligned} P(|X - E[X]| \geq \epsilon) &= P((X - E[X])^2 \geq \epsilon^2) \\ &= P(Y \geq \epsilon^2) \\ &\leq \frac{E[Y]}{\epsilon^2} \\ &= \frac{\mathrm{var}(X)}{\epsilon^2} \end{aligned} \tag{1}$$

∎

Independence and sum of random variables: Two random variables are independent independent if the knowledge of Y does not influence the results of X and vice versa. This can be expressed in terms of conditional probabilities: the (conditional) probability that Y takes a certain value, say β, does not change if we know that X takes a value, say α. In other words

$$Y \text{ is independent of } X \text{ if } P(Y=\beta|X=\alpha) = P(Y=\beta) \text{ for all } \alpha, \beta$$

But using the definition of conditional probability we find that

$$P(Y=\beta|X=\alpha) = \frac{P(Y=\beta \cap X=\alpha)}{P(X=\alpha)} = P(Y=\beta)$$

or

$$P(Y=\beta \cap X=\alpha) \;=\; P(X=\alpha)P(Y=\beta)\,.$$

This formula is symmetric in X and Y and so if Y is independent of X then X is also independent of Y and we just say that X and Y are independent.

X and Y are **independent** if $P(Y=\beta \cap X=\alpha) = P(X=\alpha)P(Y=\beta)$ for all α, β

Theorem 3. *Suppose X and Y are independent random variable. Then we have*

1. $E[XY] = E[X]E[Y]$.
2. $\mathrm{var}(X+Y) \;=\; \mathrm{var}(X) + \mathrm{var}(Y)$.

Proof. : If X and Y are independent we have

$$\begin{aligned} E[XY] &= \sum_{\alpha,\beta} \alpha\beta P(X=\alpha Y=\beta) \\ &= \sum_{\alpha,\beta} \alpha\beta P(X=\alpha)P(Y=\beta) \\ &= \sum_{\alpha} \alpha P(X=\alpha) \sum_{\beta} \beta P(Y=\beta) \\ &= E[X]E[Y] \end{aligned} \tag{2}$$

To compute the variance of $X+Y$ we note first that if X and Y are independent so are $X-E[X]$ and $Y-E[Y]$ and that $X-E[X]$ and $Y-E[Y]$ have expected value 0. We find then

$$
\begin{aligned}
\operatorname{var}(X+Y) &= E\left[((X+Y)-E[X+Y])^2\right] \\
&= E\left[((X-E[X])+(Y-E[Y]))^2\right] \\
&= E\left[(X-E[X])^2+(Y-E[Y])^2+2(X-E[X])(Y-E[Y])\right] \\
&= E\left[(X-E[X])^2\right]+\left[(Y-E[Y])^2\right]+2\left[E(X-E[X])(Y-E[Y])\right] \\
&= E\left[(X-E[X])^2\right]+\left[(Y-E[Y])^2\right]+2\left[E(X-E[X]]\,E\left[Y-E[Y]\right]\right] \\
&= E\left[(X-E[X])^2\right]+\left[(Y-E[Y])^2\right] \\
&= \operatorname{var}(X)+\operatorname{var}(Y) \qquad (3)
\end{aligned}
$$

∎

Exercise 1: What does Chebyshev say about the probability that a random variable X takes values within an interval of size k times the variance of X centered around the mean of X?

Exercise 2:

1. Let $r > 1$ be a real number. Consider the random variable X which takes values r with probability p and $1/r$ with probability $1-p$. Compute the expected value $E[X]$, $E[X^2]$ and the variance of X.

2. A very simple model for the price of a stock suggests that in any given day (independently of any other days) the price of a stock q will increase by a factor r to qr with probability p and decrease to q/r with probability $1-p$. Suppose we start with a stock of price 1. Find a formula for the mean and the variance of the price of the stock after n days. *Hint:* Use n copies of the random variable in part 1.