2 Random variables and distributions (cont'd)

2.5 Expectation (cont'd)

$$\exists(g(x)) = \begin{cases} \sum_{i=1}^{\infty} g(x_i)\mathbb{P}(X = x_i) & \text{for discrete } X \\ \int_{-\infty}^{\infty} g(x)f(x)dx & \text{for continuous} X \end{cases}$$

Properties of expectation

1. Linearity:expectation of X: $\mathbb{E}(X) = \begin{cases} \sum X_i \mathbb{P}(X = x_i) \\ \int_{-\infty}^{x_1} xf(x)dx \end{cases}$, $g(X) = x$
   - $\mathbb{E}(ax + b) = a\mathbb{E}(x) + b$
   - $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
2. If $X \perp\!\!\!\perp Y$, then $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X)) \cdot \mathbb{E}(h(Y))$
   - **proof**: (continuous case)
   - $$\begin{aligned} \mathbb{E}(g(X)h(Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(f)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} g(x)f_X(x) \cdot \int_{-\infty}^{\infty} h(y)f_Y(y)dy \end{aligned}$$
   - In particular, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ if $X \perp\!\!\!\perp Y$

Definitions

Definition: The expectation $\mathbb{E}(X^n)$ is called the n-th moment of X:

- 1st moment: $\mathbb{E}(X)$
- 2st moment: $\mathbb{E}(X^2)$

Definition: The variance of a r.v X is defined as:

$$Var(x) = \mathbb{E}((X - \mathbb{E}(X))^2) \text{ also denoted as } \sigma^2, \sigma_x^2$$

Definition: the covariance of the r.v's X and Y is defined as:

$$Cov(X, Y) = \mathbb{E}((X - \mathbb{E}(X)))\mathbb{E}((Y - \mathbb{E}(Y)))$$

Thus $Var(X) = Cov(X, X)$

Definition: the correlation between X and Y is defined as:

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Fact: $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

Proof:

$$\begin{aligned} Var(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + (\mathbb{E}(X)^2)) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X\mathbb{E}(X)) + (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - 2(\mathbb{E}(X))^2 + (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \end{aligned}$$

Fact: $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

Proof: similar to previous

Variance and covariance are **translation invariant**. Variance is quadratic, covariance is bilinear.

$$Var(aX + b) = a \cdot Var(X)$$

$$Cov(aX + b, cY + d) = ac \cdot Cov(X, Y)$$

Proof:

$$\begin{aligned} Var(aX + b) &= \mathbb{E}((aX + b0\mathbb{E}(aX + b)^2)) \\ &= \mathbb{E}([a(X - \mathbb{E}(X))]^2) \\ &= a^2\mathbb{E}((X - \mathbb{E}(X)^2)) \\ &= a^2\mathbb{E}(X) \end{aligned}$$

Proof: $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

Exercise

If $X \perp\!\!\!\perp Y$, then $Cov(X, Y) = 0$ and $Var(X + Y) = Var(X) + Var(Y)$

Proof:

$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
we know:
$X\ Y \Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
Thus, $Cov(X,Y) = 0 \Rightarrow Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
So we see independence $\Rightarrow$ Covariance is 0: "uncorrelated"
the converse is not true.
$Cov(X,Y) = 0 \Rightarrow \not{i}$ndependence

Remarks

We have $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

If $X \perp\!\!\!\perp Y$, we also have:

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, and
- $Var(X+Y) = Var(X) + Var(Y)$

It's important to remember that the first result and the other two results are of very different nature. While $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ is a property of expectation and holds unconditionally;

the other two, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ and $Var(X+Y) = Var(X) + Var(Y)$, only hold if $X \perp\!\!\!\perp Y$.

It is more appropriate to consider them as **properties of independence** rather than properties of expectation and variance

2.6 Indicator

A random variable I is called an indicator, if

$$I(w) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \not\in A \end{cases}$$

for some event A

For A given, I is also elevated as I_A

The most important property of indicator is its expectation gives the probability of the event $\mathbb{E}(I_A) = \mathbb{P}(A)$

Proof:

$$
\begin{aligned}
\mathbb{P}(I_A=1) &= \mathbb{P}(\omega : I_A(\omega=1)) \\
&= \mathbb{P}(\omega : \omega \in A) \\
&= \mathbb{P}(A)
\end{aligned}
$$

$$\mathbb{P}(I_A=0) = 1-\mathbb{P}(A) \Rightarrow \mathbb{E}(I_A) = 1\cdot\mathbb{P}(A) + c\cdot(1-\mathbb{P}(A)) = \mathbb{P}(A)$$

Example 1: we see $I_A \sim Ber(\mathbb{P}(A))$

Let $X \sim Bin(n,p)$, X is number of successes in n Bernoulli trials, each with probability p of success

$$\Rightarrow X = I_1 + \cdots + I_n$$

where $I_1, \cdots, I_n$ are indicators for independent events. $I_i = 1$ if th i the trial is a success. $I_i = 0$ if the i th trial is a failure.

Hence I_i are **i.d.**(independent and identically distributed) r.v's

$$
\begin{aligned}
\Rightarrow \mathbb{E}(X) &= \mathbb{E}(I_1 + \cdot, I_N) \\
&= \mathbb{E}(I_1) + \cdots | \mathbb{E}(I_n) \\
&= p + \cdots + p = n\cdot p
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= Var(I_1 + \cdots + I_n) \\
&= Var(I_1) + \cdots + Var(I_n) \\
&= n\cdot Var(I_i) \\
&= n\cdot p(1-p)
\end{aligned}
$$

$$Var(I_1) = \mathbb{E}(I_1^2) - (\mathbb{E}(I_1))^2 = \mathbb{E}(I_1) - (\mathbb{E}(I_1))^2 = p - p^2 = p(1-p)$$