Module 9.3: Linearity of Expectation

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Sum of two r.v.'s

Let $X, Y : \Omega \to \mathbb{R}$ be two random variables defined on the same probability space $(\Omega, \Pr)$. Their **sum**, notation $X + Y$, is the random variable on the same space defined by $(X + Y)(w) = X(w) + Y(w)$ for all $w \in \Omega$.

In general there is no simple relation between the distribution of $X + Y$ and those of X and Y (see examples).

Example. In the green-purple dice space consider the r.v. G that takes as values the numbers shown by the green die.

G has the same distribution as D (see earlier). Same with P (purple die).

The sum $G + P$ is the same as the r.v. S (see earlier).

However, S is **not** $D + D$ because S and D are defined on **different** probability spaces.

Scalar multiplication of an r.v.

Let $c \in \mathbb{R}$ and $X : \Omega \to \mathbb{R}$ be an r.v. The **multiplication by** c of X, notation $c\,X$ is the random variable on the same space defined by $(c\,X)(w) = c \cdot X(w)$ for all $w \in \Omega$. Its distribution is closely related to that of X (unless $c = 0$).

Example. Let D be the single die r.v. we saw earlier. $2D$ takes as values $2, 4, 6, 8, 10, 12$ with the uniform distribution.

Example. We flip a fair coin n times. Let H, respectively T, be the r.v. that returns the number of heads, respectively tails, observed.

Consider H, T and $(-1)T$.

$H + T$ is quite trivial: takes one value, n, with probability 1!

But $H - T = H + (-1)T$ is very interesting: intuitively $\mathrm{E}[H - T] = 0$.

Linearity of expectation

Proposition. Let $X_1, \ldots, X_n$ be random variables on the same probability space (Ω, Pr) and let $c_1, \ldots, c_n \in \mathbb{R}$. Then:

$$\mathsf{E}[c_1 X_1 + \cdots + c_n X_n] = c_1 \mathsf{E}[X_1] + \cdots + c_n \mathsf{E}[X_n]$$

Proof.

$$\begin{aligned} \mathsf{E}\left[\sum_{i=1}^{n} c_i X_i\right] &= \sum_{w \in \Omega} \left(\sum_{i=1}^{n} c_i X_i\right)(w) \cdot \mathrm{Pr}[w] \\ &= \sum_{w \in \Omega} \left(\sum_{i=1}^{n} c_i X_i(w)\right) \cdot \mathrm{Pr}[w] \\ &= \sum_{i=1}^{n} c_i \left(\sum_{w \in \Omega} X_i(w) \cdot \mathrm{Pr}[w]\right) = \sum_{i=1}^{n} c_i \mathsf{E}[X_i] \end{aligned}$$

Applications of linearity of expectation

Problem. Compute $\mathrm{E}[S]$ where S is the r.v. that returns the sum of numbers shown by two fair dice rolled together.

Answer. Assume green-purple dice. Since G and P have the same values and distribution as D: $\mathrm{E}[G] = \mathrm{E}[P] = \mathrm{E}[D] = 3.5$. Now, using linearity of expectation, $\mathrm{E}[S] = \mathrm{E}[G+P] = \mathrm{E}[G] + \mathrm{E}[P] = 3.5 + 3.5 = 7$.

Problem. We flip a fair coin n times. Compute $\mathrm{E}[H - T]$, the difference between the number of heads and of tails observed.

Answer. H and T have the same distribution, therefore $\mathrm{E}[H] = \mathrm{E}[T]$. Then, by linearity of expectation,
$\mathrm{E}[H - T] = \mathrm{E}[H + (-1)T] = \mathrm{E}[H] + (-1)\mathrm{E}[T] = \mathrm{E}[H] - \mathrm{E}[T] = 0$.

Quiz

Recall that we tried to compute the expectation of the sum of three fair dice and gave up because the calculations were too onerous? Now, armed with **linearity of expectation**, answer the following.

We roll a fair die r times and add up the numbers shown. What is the expected value of this sum?

(A) 10.5

(B) $3.5/r$

(C) $7r/2$

Answer

We roll a fair die r times and add up the numbers shown. What is the expected value of this sum?

(A) 10.5

Incorrect. This is the expectation of the sum of **three** dice, not r.

(B) $3.5/r$

Incorrect. No reason to divide by r.

(C) $7r/2$

Correct. We add up r times the expectation of the value of a single die: $(r)(3.5) = (r)(7/2) = 7r/2$.

More information

We roll a fair die r times independently. Let D_i be the rv that returns the value shown by the i'th roll and let $W = D_1 + \ldots + D_r$ be the sum of the r values shown.

By linearity of expectation $\mathrm{E}[W] = \mathrm{E}[D_1] + \cdots + \mathrm{E}[D_r]$.

So far we did not use the fact that the rolls are independent (linearity of expectation does not require any such assumption).

But now, because the rolls are independent we can assert that the probability distribution of each D_i is the same as the probability distribution of a single die roll, D. And we have calculated earlier that $\mathrm{E}[D] = 3.5$.

Therefore $\mathrm{E}[D_i] = 3.5$ for $i = 1, \ldots, r$ thus $\mathrm{E}[W] = (r)(3.5)$.