Module 10.2: Variance

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Measuring deviation from mean

If X is an r.v. its **expectation**, $\mathrm{E}[X]$, also called its **mean**, is often denoted by $\mu = \mathrm{E}[X]$.

The **deviation** of X from its mean, $X - \mu$ is also an r.v.

However, if we take $\mathrm{E}[X - \mu]$ we get, by the linearity of expectation $\mathrm{E}[X] - \mathrm{E}[\mu] = \mu - \mu = 0$!

This is not very informative. We do not want the positive and the negative deviations from the mean to cancel each other out.

We could take the expectation of the absolute value of $X - \mu$.

However, working with absolute values is mathematically messy.

As a result, statisticians decided that squaring $X - \mu$ is more useful!

Variance

The **variance** of a random variable X is defined as

$$\mathrm{Var}[X] \;=\; \mathrm{E}[(X-\mu)^2] \qquad (\text{where } \mu = \mathrm{E}[X])$$

The **standard deviation** of a random variable X is

$$\sigma[X] \;=\; \sqrt{\mathrm{Var}[X]}$$

Why square root? By undoing the squaring in the variance we obtain the same units of measurement as those used for the values of the random variable.

Notation. When the random variable is understood , its mean is often denoted by μ, its standard deviation by σ, and its variance by σ^2.

An alternative formula for variance

Proposition. Let X be an r.v. defined on (Ω, Pr).

$$\mathsf{Var}[X] \;=\; \mathsf{E}[X^2] \;-\; \mu^2 \qquad\qquad (\text{where } \;\forall w \in \Omega \;\; X^2(w) = (X(w))^2\;)$$

Proof. Using linearity of expectation and the fact that μ is a constant:

$$\begin{aligned}
\mathsf{E}[(X-\mu)^2] \;&=\; \mathsf{E}[X^2 - 2\mu X + \mu^2] \\
&=\; \mathsf{E}[X^2] - 2\mu\mathsf{E}[X] + \mu^2 \\
&=\; \mathsf{E}[X^2] - 2\mu^2 + \mu^2 \;=\; \mathsf{E}[X^2] - \mu^2
\end{aligned}$$

Penn Engineering

Activity : Variance of a fair die

In the lecture segment "Random variables" we introduced the random variable D that returns the number shown by a fair die. This is a uniform r.v. with $\mathrm{Val}(D) = [1..6]$. Later, in the lecture segment "Expectation" we calculated $\mathrm{E}[D] = 3.5$. In this activity we will calculate $\mathrm{Var}[D]$, the **variance** of the number shown by a fair die.

Question. What is the distribution of the r.v. D^2?

In the video, there is a box here for learners to put in an answer to the question above. As you read these notes, try it yourself using pen and paper!

ACTIVITY : Variance of a fair die (continued)

Answer. Like D, the r.v. D^2 is **uniform**. We have $\text{Val}(D^2) = \{1, 4, 9, 16, 25, 36\}$ and the probability is $1/6$ for each of these.

This is a particular case of a more general fact: If X is an r.v. that takes positive values then the distribution of X^2 is very similar to that of X, with the same probabilities, but the values taken are squared.

Activity : Variance of a fair die (continued)

Now to the calculation of Var[D]. Using the alternative formula for variance we have

$$\begin{aligned}\mathrm{Var}[D] &= \mathrm{E}[D^2] - \mathrm{E}[D]^2 \\ &= \frac{1}{6}(1+4+9+16+25+36) - (3.5)^2 \\ &= \frac{35}{12} \simeq 2.92\end{aligned}$$