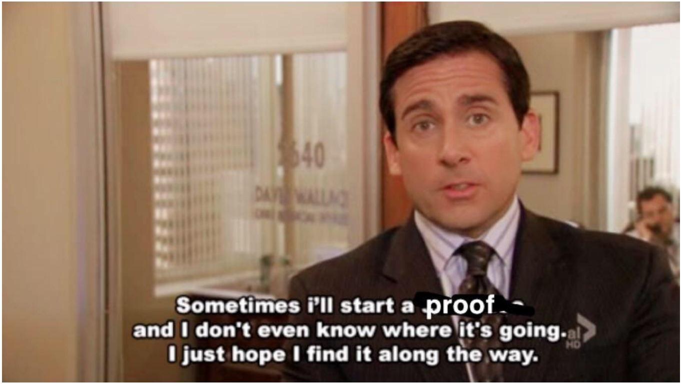
STAT 230 SECTION 2 LECTURE 19



Today's Agenda

Last time:

- Expected value

Today (Lec 19, 06/15):

- Expected value of Binomial and Poisson
- Definition of variance

Chapter 7: Expected Value and Variance

Definition

Suppose X is a discrete random variable with probability function $f(x)$. Then $E(X)$ is called the **expected value** of X, and is defined by

$$E(X) = \sum_{x \in X(S)} x\, f(x).$$

The expected value of X is sometimes referred to as the **mean**, **expectation**, or the **first moment** of X.

Properties

- If $g : \mathbb{R} \to \mathbb{R}$ is a function, then

$$E(g(X)) = \sum_{x \in X(S)} g(x) f(x).$$

  "Law of unconscious statistician"

- For constants $a, b \in \mathbb{R}$,

$$E(aX + b) = aE(X) + b$$

  "Linearity of the expectation"

Question

Suppose the discrete random variable X has probability function

$$f(-1) = 0.15, \quad f(0) = 0.2, \quad f(1) = 0.5, \quad f(2) = 0.05, \quad f(3) = 0.1.$$

Compute $E(X)$ and $E(\exp(X))$.

Going forward, we would like to compute $E[W]$, $E[X]$, $E[Y]$, and $E[Z]$ if

a) $W \sim Binomial(n, p)$

b) $X \sim hyp(N, r, n)$

c) $Y \sim NB(k, p)$

d) $Z \sim Poi(\mu)$

Mean of Binomial

If $W \sim Binomial(n, p)$, then $E[W] = np$.

Question

Suppose two fair six sided die are independently rolled 24 times, and let X denote the number of times the sum of die rolls is 7.

a) $E(X) = 2$
b) $E(X) = 3$
c) $E(X) = 4$
d) $E(X) = 5$
e) $E(X) = 6$
f) $E(X) = 7$
g) $E(X) = 8$

Mean of Poisson

If $Z \sim Poi(\mu)$, then $E[Z] = \mu$.

Example

Suppose that calls to the Canadian Tire Financial call center follow a Poisson process with rate 30 calls per minute. Let X denote the number of calls to the center after 1 hour. Compute $E[X]$.

Using similar techniques, we can show the following.

If $X \sim hyp(N, r, n)$, then

$$E[X] = n\frac{r}{N}.$$

If $Y \sim NB(k, p)$, then

$$E[Y] = \frac{k(1-p)}{p}.$$

With the expectation, we now know about the weighted average value of the distribution.

On that note, think about this:

- X is a r.v. representing the outcome of one fair 6-sided die roll
- Y is a r.v. representing the number of phone calls over 1 minute at Lenovo call centre, with the rate of 3.5 calls per minute

Are they similar in any ways? How are they different?

- X and Y have the same expectation:
  $\mu_X = \mathrm{E}[X] = 3.5 = \mathrm{E}[Y] = \mu_Y$.

- BUT the extent of deviation from the expectations are different:
  - For X, 1 or 6 are just as likely to appear as 2 or 5.
  - For Y, values closer to 3definition.5 are more likely to appear, but you could possibly get an insanely high number too. (ex: order backlog from Black Friday)

Measure deviation from the mean

As seen in the example, expectation alone may not be enough often times; we may want to study how much the random variable tends to deviate from its mean. Let $\mathsf{E}[X] = \mu$.

a) Deviation

$$\mathsf{E}[(X - \mu)] = \mathsf{E}[X] - \mu = 0$$

b) Absolute deviation

$$\mathsf{E}\left[|X - \mu|\right]$$

c) Squared deviation

$$\mathsf{E}\left[(X - \mu)^2\right]$$

The squared deviation turns out to be a particularly useful measure of variability.

Definition

The **variance** of a random variable X is denoted $Var(X)$, and is defined by

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

Remark:

- Note that $Var(X) \geq 0$.
- A simple calculation gives the "short cut formula" to compute $Var(X)$:
$$Var(X) = E(X^2) - (E(X))^2.$$
- Similarly,

$$Var(X) = E(X(X-1)) + E(X) - (E(X))^2.$$

Example

Let X denote the outcome of a fair six sided die roll. Compute $Var(X)$.