Aaaaand back to campus. On a Monday.

Today's Agenda

Last time:

- Chapter 5!

Note the R Tutorial on Learn!

Today (Lec 18, 06/13):

- Start of Chapter 7!

Chapter 7: Expected Value and Variance

Some key concepts in this chapter are:

- Expected value
- Mean of a distribution
- Variance of a distribution

Probability and statistics are closely related to data; we often try to "extract" additional information from data.

One of the simplest ways to do this organising the data into a **frequency table**.

Number of Cups of Coffee	Tally	Frequency
0 - 3	//	2
4 - 7	///	3
8 - 11	~~////~~ ///	8
12 - 15	///	3
16 - 19	//	2

Figure: I'm one of the two in the 16-19 cups slot, FYI.

We can take the frequency table to draw various graphs.

However, sometimes, a numerical summary (we call them "statistics") can provide desired information more concisely.

How do you compute the average grade for a test?

A Add everyone's marks and divide by the number of grades you added

B Ask a smart friend

C Google it

D Check on LEARN before the instructor takes it down

Definition (Sample mean)

Let $x_1,\ x_2, \ldots,\ x_n$ be n outcomes for a random variable X (such a set is called a **sample**). Then, its **sample mean** is defined as

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

While the sample mean is the most common summary statistic, other important statistics include

- **median**: a value such that half of the results are below it and the other half above it, when the sample is arranged in numerical order.
- **mode**: The most frequently-occuring value in a sample. NOTE: a sample may have more than one mode.

We can do something similar to a random variable X directly if we know its distribution. Such a mean would be "theoretical", as we are working from its distribution rather than an actual sample.

Definition

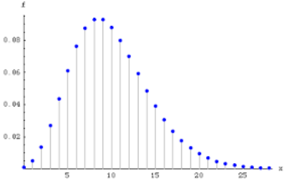
Suppose X is a discrete random variable with probability function $f(x)$. Then $E(X)$ is called the **expected value** of X, and is defined by

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

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

Interpretations of the expected value: $E(X)$

a) $E(X)$ is the "balancing point" of the probability function $f_X(x)$



b) $E(X)$ is what the average of many, many independent realizations of the random variable X would approach (Law of large numbers).

Example

Suppose that X denotes the outcome of one fair six sided die roll. Compute $E(X)$.

Example

A lottery is conducted in which 7 numbers are drawn without replacement between the numbers 1 and 50. A player wins the lottery if the numbers selected on their ticket match all 7 of the drawn numbers. A ticket to play the lottery costs \$1, and the jackpot is valued at \$5,000,000. Compute the expected return for this bet.

Question

Suppose X is a random variable satisfying $a \leq X(\omega) \leq b$ for all $\omega \in S$. Which of the following is false:

A It is possible that $E[X] > b$

B It is possible that $E[X] > a$

C It is possible that I can get a jelly if I choose C

Remark

If

$$g : \mathbb{R} \to \mathbb{R},$$

then for a random variable X with probability function f, $g(X)$ is a random variable taking values $g(X(S))$. Therefore,

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

Example

If $g(x) = x^2$, and X is the result of a fair six sided die roll, then compute $E[g(X)]$.

Remark

If $g(x)$ is a linear function $g(x) = ax + b$, then for a random variable X

$$E[aX + b] = aE[X] + b$$

In general, $g(E[X]) \neq E[g(X)]$, although this is a common mistake.

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

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 $E(X) = 4$

B $E(X) = 5$

C $E(X) = 6$

D $E(X) = 7$

E $E(X) = 8$

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 $E[X]$.