Expectations

Mathematical Biostatistics Boot Camp

Brian Caffo, PhD
Johns Hopkins Bloomberg School of Public Health

Table of contents

Expected values

- The {\bf expected value} or {\bf mean} of a random variable is the center of its distribution
- For discrete random variable X with PMF $p(x)$, it is defined as follows

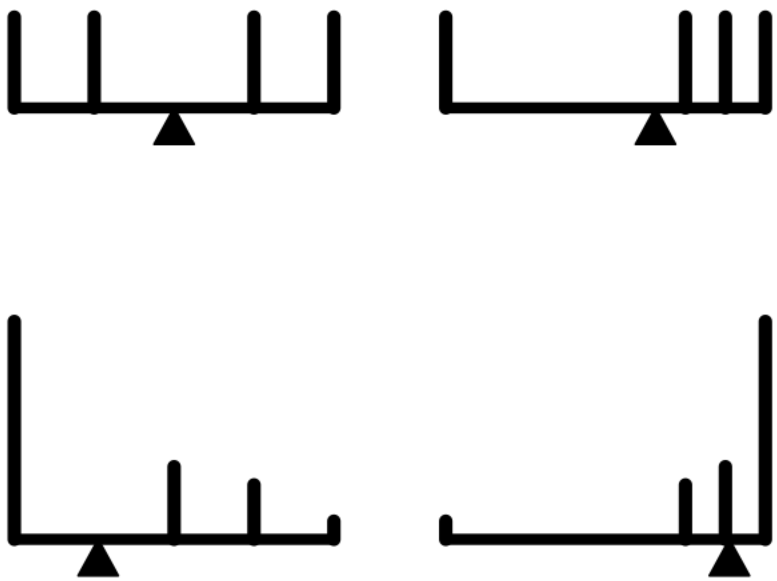
$$E[X] = \sum_x xp(x).$$

  where the sum is taken over the possible values of x

- $E[X]$ represents the center of mass of a collection of locations and weights, $\{x, p(x)\}$

3/21

Example



Example

- Suppose a coin is flipped and X is declared 0 or 1 corresponding to a head or a tail, respectively
- What is the expected value of X?

$$E[X] = .5 \times 0 + .5 \times 1 = .5$$

- Note, if thought about geometrically, this answer is obvious; if two equal weights are spaced at 0 and 1, the center of mass will be $.5$

Example

- Suppose that a die is rolled and X is the number face up
- What is the expected value of X?

$$E[X] = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5$$

- Again, the geometric argument makes this answer obvious without calculation.

Continuous random variables

- For a continuous random variable, X, with density, f, the expected value is defined as follows

$$E[X] = \int_{-\infty}^{\infty} tf(t)dt$$

- This definition borrows from the definition of center of mass for a continuous body

Example

- Consider a density where $f(x) = 1$ for x between zero and one
- (Is this a valid density?)
- Suppose that X follows this density; what is its expected value?

$$E[X] = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = 1/2$$

Rules about expected values

- The expected value is a linear operator
- If a and b are not random and X and Y are two random variables then
  - $E[aX + b] = aE[X] + b$
  - $E[X + Y] = E[X] + E[Y]$
- *In general* if g is a function that is not linear,

$$E[g(X)] \neq g(E[X])$$

- For example, in general, $E[X^2] \neq E[X]^2$

9/21

Example

- You flip a coin, X and simulate a uniform random number Y, what is the expected value of their sum?

$$E[X+Y] = E[X] + E[Y] = .5 + .5 = 1$$

- Another example, you roll a die twice. What is the expected value of the average?
- Let X_1 and X_2 be the results of the two rolls

$$E[(X_1 + X_2)/2] = \frac{1}{2}(E[X_1] + E[X_2]) = \frac{1}{2}(3.5 + 3.5) = 3.5$$

Example

1. Let X_i for $i = 1, \ldots, n$ be a collection of random variables, each from a distribution with mean μ
2. Calculate the expected value of the sample average of the X_i

$$
\begin{aligned}
E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] &= \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E[X_i] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu = \mu.
\end{aligned}
$$

11/21

Remark

- Therefore, the expected value of the **sample mean** is the population mean that it's trying to estimate
- When the expected value of an estimator is what its trying to estimate, we say that the estimator is **unbiased**

The variance

- The variance of a random variable is a measure of {\em spread}
- If X is a random variable with mean μ, the variance of X is defined as

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

the expected (squared) distance from the mean

- Densities with a higher variance are more spread out than densities with a lower variance

- Convenient computational form

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

- If a is constant then $Var(aX) = a^2 Var(X)$
- The square root of the variance is called the **standard deviation**
- The standard deviation has the same units as X

Example

- What's the sample variance from the result of a toss of a die?
    - $E[X] = 3.5$
    - $E[X^2] = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = 15.17$
- $Var(X) = E[X^2] - E[X]^2 \approx 2.92$

Example

- What's the sample variance from the result of the toss of a coin with probability of heads (1) of p?
  - $E[X] = 0 \times (1 - p) + 1 \times p = p$
  - $E[X^2] = E[X] = p$
- $Var(X) = E[X^2] - E[X]^2 = p - p^2 = p(1 - p)$

16/21

Example

- Suppose that a random variable is such that $0 \leq X \leq 1$ and $E[X] = p$
- Note $X^2 \leq X$ so that $E[X^2] \leq E[X] = p$
- $Var(X) = E[X^2] - E[X]^2 \leq E[X] - E[X]^2 = p(1-p)$
- Therefore the Bernoulli variance is the largest possible for random variables bounded between 0 and 1

Interpreting variances

- Chebyshev's inequality is useful for interpreting variances
- This inequality states that

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

- For example, the probability that a random variable lies beyond k standard deviations from its mean is less than $1/k^2$

$$\begin{aligned} 2\sigma &\rightarrow 25\% \\ 3\sigma &\rightarrow 11\% \\ 4\sigma &\rightarrow 6\% \end{aligned}$$

- Note this is only a bound; the actual probability might be quite a bit smaller

Proof of Chebyshev's inequality

$$
\begin{aligned}
P(|X-\mu| > k\sigma) &= \int_{\{x:|x-\mu|>k\sigma\}} f(x)dx \\
&\leq \int_{\{x:|x-\mu|>k\sigma\}} \frac{(x-\mu)^2}{k^2\sigma^2} f(x)dx \\
&\leq \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{k^2\sigma^2} f(x)dx \\
&= \frac{1}{k^2}
\end{aligned}
$$

Example

- IQs are often said to be distributed with a mean of 100 and a sd of 15
- What is the probability of a randomly drawn person having an IQ higher than 160 or below 40?
- Thus we want to know the probability of a person being more than 4 standard deviations from the mean
- Thus Chebyshev's inequality suggests that this will be no larger than 6\%
- IQs distributions are often cited as being bell shaped, in which case this bound is very conservative
- The probability of a random draw from a bell curve being 4 standard deviations from the mean is on the order of 10^{-5} (one thousandth of one percent)

Example

- A popular buzz phrase in industrial quality control is Motorola's``Six Sigma'' whereby businesses are suggested to control extreme events or rare defective parts
- Chebyshev's inequality states that the probability of a ``Six Sigma'' event is less than $1/6^2 \approx 3\%$
- If a bell curve is assumed, the probability of a ``six sigma'' event is on the order of 10^{-9} (one ten millionth of a percent)