The Expected Value of a Random Variable

When observing a random variable, say *X*, with a binomial probability distribution, we can't say what value *X* will take. The possible values are {0,1,2,...,n} and the probability that *X* will take some value k is

$$P(X=k)=\binom{n}{k}\theta^{k}(1-\theta)^{n-k} \quad k=0,1,2,...,n$$

⇨ *The expected value of X is defined to be the weighted average of all the possible values, where the weight assigned to each value is its probability:*

$$E(X) = 0P(X=0) + 1P(X=1) + 2P(X=2) + ... + nP(X=n)$$

$$= (0\, w_0) + (1\, w_1) + (2\, w_2) + ...+ (n\, w_n)$$

A value that is more probable gets more weight, and one that is very improbable gets very little weight.

E(X) is "the expected value of *X*," or "the expected number of successes." For a binomial random variable, the expected value is

$$E(X) = n\theta$$

- If $\theta = 1$ (success is certain on every trial), then the expected number of successes is the number of trials, $E(X) = n$.
- If $\theta = 0$ (success is impossible), then the expected number of successes is zero, $E(X) = 0$.
- If $\theta = 1/2$ (success and failure are equally likely), then the expected number of successes is $E(X) = n/2$, half the number of trials.
- If $\theta = 1/4$ (probability of success is one fourth), then the expected number of successes is $E(X) = n/4$, one fourth of the number of trials.

The <u>proportion</u> of successes, X/n, is also a random variable.

Because $P(X/n = k/n)$ is the same as $P(X=k)$, we can find the expected value of the proportion of successes:

$$
\begin{aligned}
E(X/n) &= (0/n)\,P(X/n = 0/n) + (1/n)\,P(X/n = 1/n) + \ldots \\
&\qquad + (n/n)\,P(X/n = n/n) \\
&= (0/n)\,P(X=0) + (1/n)\,P(X=1) + \ldots + (n/n)\,P(X=n) \\
&= (1/n)\,[0\,P(X=0) + 1\,P(X=1) + \ldots + n\,P(X=n)] \\
&= (1/n)\,E(X) \\
&= (1/n)\,n\theta \\
&= \theta
\end{aligned}
$$

The expected value of the sample proportion, X/n, of successes is simply the <u>probability</u> of success.

When the value of θ is unknown (as in the case of my 40¢ coin), we might use the proportion of successes in a sample as an estimate of θ.

⇨ When the expected value of an estimate equals the quantity that is being estimated, the estimate is said to be **unbiased**.

Thus since **E(*X*/n) = θ**, the proportion of successes in a sample, *X*/n, is an unbiased estimate of the probability of success, θ

Note: the expected value of *X* is not the same thing as the most probable value of *X*. That is, E(*X*) is not the value of *X* that has the greatest probability of occurrence.

In fact, it often happens that the expected value of *X* is not even one of the possible values of *X*.

For example, this happens whenever nθ is not an integer. Thus when n = 2, the number of successes must be 0, 1, or 2, but if θ = 1/3, the expected number of successes is 2/3.

⇨ The expected value of *X*, is also called "the mean of *X* ", "the mean of the binomial(n,θ) distribution", and the "population mean."

If *X* has a Poisson(λ) distribution, the expected value of *X* is

$$E(X) = 0\,P(X = 0) + 1\,P(X = 1) + 2\,P(X = 2) + \ldots.$$

This sum goes on forever, because there is no limit to how large *X* can be. But we can write it another way, and when we do, we can find its value:

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \; P(X = k) \\
&= \sum_{k=0}^{\infty} k \; \lambda^{k} e^{-\lambda} / k! \\
&= \sum_{k=1}^{\infty} k\lambda \; \lambda^{(k-1)} e^{-\lambda} / k(k-1)! \\
&= \lambda \sum_{k=1}^{\infty} \lambda^{(k-1)} e^{-\lambda} / (k-1)! \\
&= \lambda \sum_{i=0}^{\infty} \lambda^{i} e^{-\lambda} / i! \\
&= \lambda \sum_{i=0}^{\infty} P(X = i) \\
&= \lambda
\end{aligned}
$$

So E(X)=λ for a Poisson distribution.

Not all random variables <u>have</u> expected values. Binomial, Hypergeometric, and Poisson random variables <u>do</u> have expected values, as we have already seen. But this is not true of all probability distributions. We will see later how it can happen that a random variable does not have an expected value. For now, we just note that it <u>can</u> happen.

Here are the expected values for the distributions we have covered so far:

Distribution	**E(X)**
Bin(n, θ)	nθ
Poiss(λ)	λ
Hyper(N,n,C)	Cn/N
N(μ,σ^2)	μ
Unifrom(a,b)	(b+a)/2

More on Expected Values

The expected value of a discrete-valued random variable, *X*, is a weighted average of the all the possible values, the weight at each value being the probability of that value.

When the set of possible values is S = {0,1,2, ...,n}, the expected value is

$$E(X) = \sum_{k=0}^{n} k \quad P(X = k)$$

For any discrete-valued random variable *X*

$$E(X) = \sum_{all\ possible\ values} (value)\ P(X = value)$$

For a random variable with a continuous probability distribution the expected value is found by integration:

$$E(X) = \int_{-\infty}^{\infty} x \quad f(x) \ dx$$

If we add a constant, say 3, to a random variable X, we get a new random variable, $X + 3$, whose expected value is

$$E(X + 3) = E(X) + 3.$$

Adding 3 to X increases its value by three, so it also increases its expected value by 3. (Subtracting 3 decreases the expected value by 3.) This is true for <u>any</u> constant, c, positive or negative:

$$E(X + c) = E(X) + c.$$

Similarly, if we multiply X by 3, we get a random variable, $3X$, whose expected value is three times the expected value of X:

$$E(3X) = 3E(X).$$

And, in general, if we multiply by <u>any</u> constant, c,

$$E(cX) = cE(X).$$

⇨⇨ **For <u>any</u> constants, a and b** **

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

<u>Example</u>: If X ~ binomial(n,θ), then

(a) $X-n\theta$ has expected value zero.

$$E(X-n\theta) = E(X)-n\theta = n\theta - n\theta = 0.$$

(b) X/n has expected value θ.

$$E(X/n) = E(X)/n = n\theta/n = \theta.$$

(c) $X/n - \theta$ has expected value zero.

$$E(X/n - \theta) = E(X/n) - \theta = \theta - \theta = 0.$$

Functions of a random variable *X* are themselves random variables with their own expected value.

This is true, not only for the simple functions like a*X*+b that we just looked at, but of more complicated functions as well.

When *X* has a discrete probability distribution, the expected value of any function of *X*, say g(*X*), is

$$E(\,g(X)\,) = \sum_{all\ possible values} g(value)P(\,X = value\,)$$

When *X* has a continuous distribution, the expected value is

$$E(X\ \ ) = \int_{-\infty}^{\infty} g(x)\ f\ (x)\,dx$$

Example

If X ~ binomial(n,θ), what is the expected value of g(*X*) = X^2 ?

$$E(X^2) = 0^2 P(X=0) + 1^2 P(X=1) + 2^2 P(X=2) + 3^2 P(X=3) + \ldots$$

$$= \sum_{k=0}^{n} k^2 P(X=k)$$

$$= \sum_{k=0}^{n} k^2 \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

$$= n\theta(1-\theta) + n^2\theta^2$$

You do not need to know this particular result, just to be aware that it can be worked out. The same comment applies to the next one as well.

When X has a normal(μ,σ^2) probability distribution the expected value of X^2 is

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \sigma^2 + \mu^2$$

Variance (and Standard Deviation) of a Random Variable

The <u>variance</u> of a random variable, *X*, is defined as

$$\text{Var}(X)=E[\ (X-E(X))^2\ ]$$

That is, the variance is the expected value of the squared distance between *X* and its mean, E(*X*).

The <u>standard</u> <u>deviation</u> of *X* is the square root of its variance,

$$SD(X)=\sqrt{Var(X)}$$

Example:

What is the variance of a Bernoulli(θ) random variable?

The possible values of *X* are 0 and 1, and we know that E(X) is θ, so *X*-E(*X*) = *X*-θ.

Thus the possible values of $(X\text{-}E(X))^2$ are $(0\text{-}\theta)^2$ (which is just θ^2) and $(1\text{-}\theta)^2$. These values have probabilities of (1-θ) and θ.

So the variance is

$$E((X-\theta)^2) = (0-\theta)^2 P(X=0) + (1-\theta)^2 P(X=1)$$

$$= \theta^2(1-\theta) + (1-\theta)^2\theta$$

$$= \theta(1-\theta)(\theta+1-\theta)$$

$$= \theta(1-\theta)$$

The variance of a Bernoulli random variable is θ(1-θ) and the standard deviation is $\sqrt{\theta(1-\theta)}$

Here are the some expected values and variances

Distribution	E(X)	Var(X)
Bin(n, θ)	nθ	nθ(1-θ)
Poiss(λ)	λ	λ
Hyper(N,n,C)	Cn/N	Yuck
$N(\mu, \sigma^2)$	μ	σ^2
Unifrom(a,b)	(b+a)/2	$(b-a)^2/12$

Where yuck = $C\frac{n}{N}\frac{(N-n)(N-C)}{N(N-1)}$

Just like expected values, variances have rules for calculating the variance of a function of a random variable.

If we add a constant to X we increase the expected value by that amount, but we do not change the variance

$$E(X + c) = E(X) + c$$

$$Var(X + c) = Var(X)$$

If we multiply X by a constant, c, we multiply the expected value by that same constant,

$$E(cX) = cE(X),$$

and we multiply the variance by the <u>square</u> of that constant:

$$Var(cX) = c^2 Var(X).$$

Therefore multiplying X by a constant, c, multiplies the standard deviation by |c|:

$$SD(cX) = \sqrt{Var(cX)} = \sqrt{c^2 Var(X)} = |c|\sqrt{Var(X)} = |c|\,SD(X)$$

⇨ **For any constants a and b, and any random variable X**

$$\begin{aligned} E(aX + b) &= aE(X) + b \\ Var(aX + b) &= a^2\, Var(X) \\ SD(aX + b) &= |a|\, SD(X) \end{aligned}$$

<u>Examples</u>

$$\begin{aligned} E(3X - 2) &= 3\,E(X) - 2 \\ Var(3X - 2) &= 9\,Var(X) \\ SD(3X - 2) &= 3\,SD(X) \end{aligned}$$

$$\begin{aligned} E(-3X - 2) &= -3\,E(X) - 2 \\ Var(-3X - 2) &= 9\,Var(X) \\ SD(-3X - 2) &= 3\,SD(X) \end{aligned}$$

$$Var(X) = Var(-X)$$

⇨ Variances and standard deviations can <u>never</u> be negative.

Variances and standard deviations can be zero.

If Var(X) = 0, it means that X is just some fixed constant. That is, there must be some constant, c, for which P(X = c) = 1, with no random variability at all.

Example:

$X \sim$ Binomial (n,θ) so $E(X) = n\theta$, $Var(X) = n\theta(1-\theta)$.

- Var(X) is zero when $\theta = 0$, in which case failure is certain, so X can only be zero. P($X = 0$) = 1
- Var(X) is also zero when $\theta = 1$, and in that case, success is certain, so X can only be n. P(X=n)=1

In cases like these, where all of the probability is concentrated on one point, X is called a degenerate random variable. (There is really no randomness or uncertainty about what value X will have.)