Expected Value Theorems

A random variable is a variable whose value is a numerical outcome of a random process. Let X be a random variable whose possible values are $x_1, x_2, \ldots, x_N$ occur with probabilities $p(x_1), p(x_2), \ldots, p(x_N)$

Definition 1: The expected value of the discrete random variable *X*, denoted $E(X)$, is defined to be $E(X) = \sum_i x_i p(x_i)$. This is another way to define the mean of the population, so $E(X) = \mu$.

Definition 2: The variance of the discrete random variable *X*, denoted $V(X)$, is defined to be $V(X) = E\left((X-\mu)^2\right)$.

Expected Values of Linear Transformations

Theorem 1: $E(aX) = a\,E(X)$

Proof: $E(aX) = \sum_i a x_i p(x_i) = a \sum_i x_i p(x_i) = aE(X)$.

Theorem 2: $E(X+b) = E(X) + b$

Proof: $E(X+b) = \sum_i (x_i + b) p(x_i) = \sum_i x_i p(x_i) + \sum_i b p(x_i)$

$$= \sum_i x_i p(x_i) + b \sum_i p(x_i) = E(X) + b, \text{ since } \sum_i p(x_i) = 1.$$

Expected Values of Sums and Products

Theorem 3: $E(X+Y) = E(X) + E(Y)$

Proof: $E(X+Y) = \sum_i \sum_j (x_i + y_j)\, p(x_i)\, p(y_j)$

$$= \sum_i \sum_j x_i p(x_i)\, p(y_j) + \sum_i \sum_j y_j p(x_i)\, p(y_j)$$

but $x_i p(x_i)$ is constant in the summation over *j*, and $y_j p(y_j)$ is a constant in the summation over *i*. So we have

$$\sum_i x_i p(x_i) \sum_j p(y_j) + \sum_i p(x_i) \sum_j y_j p(y_j) = E(X) \cdot 1 + 1 \cdot E(Y) = E(X) + E(Y).$$

Theorem 4: $E(XY) = E(X)E(Y)$, when *X* and *Y* are independent random variables.
Proof: Same as Theorem 3.

Linear Transformations and Variance

Theorem 5: $V(X) = E(X^2) - \left(E(X)^2\right)$

Proof:
$$V(X) = E\left((X-\mu)^2\right) = E\left(X^2 - 2X\mu + \mu^2\right) = E\left(X^2\right) - 2\mu E(X) + E\left(\mu^2\right)$$
$$= E\left(X^2\right) - 2\mu^2 + \mu^2 = E\left(X^2\right) - \mu^2 = E\left(X^2\right) - \left(E(X)^2\right).$$

Theorem 6: $V(aX) = a^2 V(X)$

Proof:
$$V(aX) = E\left(a^2X^2\right) - \left(E(aX)\right)^2 = a^2E\left(X^2\right) - \left(aE(X)\right)^2$$
$$= a^2E\left(X^2\right) - a^2\left(E(X)\right)^2 = a^2V(X).$$

Theorem 7: $V(X+b) = V(X)$

Proof:
$$V(X+b) = E(X+b)^2 - \left(E(X+b)\right)^2 = E\left(X^2 + 2Xb + b^2\right) - \left(E(X)+b\right)^2$$
$$= E\left(X^2 + 2Xb + b^2\right) - \left[\left(E(X)\right)^2 + 2bE(X) + b^2\right] = V(X).$$

Variance of Sums and Differences

Theorem 8: $V(X+Y) = V(X) + V(Y)$, when the random variables *X* and *Y* are independent.

Proof:
$$V(X+Y) = E(X+Y)^2 - \left[E(X+Y)\right]^2$$
$$= E\left(X^2 + 2XY + Y^2\right) - \left[\left(E(X)\right)^2 + 2E(X)E(Y) + \left(E(Y)\right)^2\right]$$
$$= E\left(X^2\right) + E\left(Y^2\right) - \left(E(X)\right)^2 - \left(E(Y)\right)^2 = V(X) + V(Y).$$

So, $\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$.

Theorem 9: $V(X-Y) = V(X) + V(Y)$, when the random variables *X* and *Y* are independent.
Proof: Same as Theorem 8.

Suppose we generate *n* independent values of a random variable *X*. What is the expected value and variance of the sum?

Theorem 10: $E\left(X_1+X_2+X_3+\cdots+X_n\right)=nE\left(X\right)$

Proof: $E\left(X_1+X_2+\cdots+X_n\right)=E\left(X_1\right)+E\left(X_2+\cdots+X_n\right)$

$$=E\left(X_1\right)+E\left(X_2\right)+E\left(X_3+\cdots+X_n\right)=E\left(X_1\right)+E\left(X_2\right)+\cdots+E\left(X_n\right)=nE\left(X\right).$$

Theorem 11: $V\left(X_1+X_2+X_3+\cdots+X_n\right)=nV\left(X\right)$ (recall that all X_i are independent)

Proof: $V\left(X_1+X_2+\cdots+X_n\right)=V\left(X_1\right)+V\left(X_2+\cdots+X_n\right)$

$$=V\left(X_1\right)+V\left(X_2\right)+V\left(X_3+\cdots+X_n\right)=V\left(X_1\right)+V\left(X_2\right)+\cdots+V\left(X_n\right)=nV\left(X\right).$$

So, $\sigma_{X_1+X_2+\cdots X_n}=\sigma_X\sqrt{n}$.

Sampling Distribution of the Mean

What is the mean and standard deviation of the sampling distribution of the mean?

Theorem 12: $E\left(\bar{X}\right)=\mu$

Proof: $E\left(\bar{X}\right)=E\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)=\frac{1}{n}E\left(X_1+X_2+\cdots+X_n\right)=\left(\frac{1}{n}\right)\left(nE\left(X\right)\right)=\mu$.

Theorem 13: $V\left(\bar{X}\right)=\frac{\sigma_X^2}{n}$

Proof: $V\left(\bar{X}\right)=V\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)=\frac{1}{n^2}V\left(X_1+X_2+\cdots+X_n\right)=\left(\frac{1}{n^2}\right)\left(nV\left(X\right)\right)=\frac{\sigma_X^2}{n}$.

So $\sigma_{\bar{X}}=\frac{\sigma_X}{\sqrt{n}}$.

Sample Variance

Why is the denominator of the sample variance $n-1$?

Theorem 14: $E\left(\frac{\sum_i\left(X_i-\bar{X}\right)^2}{n-1}\right)=\sigma_X^2$. We use $n-1$ because the value of $\frac{\sum_i\left(X_i-\bar{X}\right)^2}{n-1}$ is an unbiased estimator of the population variance σ_X^2.

Proof: We know that $\sigma^2 = E(X-\mu)^2$. Consider $E\left(\sum_i (X_i - \bar{X})^2\right)$.

$$E\left(\sum_i (X_i - \bar{X})^2\right) = E\left(\sum_i \left[(X_i - \mu) + (\mu - \bar{X})\right]^2\right)$$

$$= \sum_i E\left((X_i - \mu)^2\right) + 2E\left(\sum_i (X_i - \mu)(\mu - \bar{X})\right) + \sum_i E(\mu - \bar{X})^2 .$$

Since $= \sum_i E\left((X_i - \mu)^2\right) = n\sigma_X^2$ and $\sum_i E(\mu - \bar{X})^2 = n\sigma_{\bar{X}}^2 = \dfrac{n\sigma_X^2}{n} = \sigma_X^2$, we have

$$= (n+1)\sigma_X^2 + 2E\left(\sum_i X_i \mu\right) - 2E\left(\sum_i X_i \bar{X}\right) - 2E\left(\sum_i \mu^2\right) + 2E\left(\sum_i \mu \bar{X}\right).$$

Now, $2E\left(\sum_i X_i \mu\right) = 2\mu \sum_i E(X_i) = 2n\mu^2$, $2E\left(\sum_i \mu \bar{X}\right) = 2\mu\left(\sum_i E(\bar{X})\right) = 2n\mu^2$,
$2E\left(\sum_i \mu^2\right) = 2n\mu^2$, and $2E\left(\sum_i X_i \bar{X}\right) = 2E\left(\bar{X}\sum_i X_i\right)$ but $\sum_i X_i = n\bar{X}$, so
$2E\left(\bar{X}\sum_i X_i\right) = 2E\left(n\bar{X}\cdot\bar{X}\right) = 2nE\left(\bar{X}^2\right)$.

So, $E\left(\sum_i (X_i - \bar{X})^2\right) = (n-1)\sigma_X^2 + 2n\mu^2 - 2nE\left(\bar{X}^2\right) - 2n\mu^2 + 2n\mu^2$. Simplifying we have
$(n+1)\sigma_X^2 + 2n\left(\mu^2 - E\left(\bar{X}^2\right)\right)$, but $\mu^2 = \left(E(\bar{X})\right)^2$, so $\left(\mu^2 - E\left(\bar{X}^2\right)\right) = -V(\bar{X})$

$$E\left(\sum_i (X_i - \bar{X})^2\right) = (n+1)\sigma_X^2 - 2n\left(V(\bar{X})\right) = (n+1)\sigma_X^2 - 2n\left(\frac{\sigma_X^2}{n}\right) = (n-1)\sigma_X^2 .$$

So, we have $E\left(\sum_i (X_i - \bar{X})^2\right) = (n-1)\sigma_X^2$. Consequently, $E\left(\dfrac{\sum_i (X_i - \bar{X})^2}{n-1}\right) = \sigma_X^2$ and $(n-1)$ is exactly the right value to produce an unbiased estimate of the population variance from a sample of *n* independent values of a random variable.

Reference:

Goldberg, Samuel, *Probability: An Introduction*, Dover Publications, New York, 1960.