The expected value of a random variable is defined as follows:

Discrete Random Variable:
\[ E[X] = \sum_{x} x P(X = x) \]

Continuous Random Variable:
\[ E[X] = \int_{all \ x} x P(X = x) \, dx \]

The expected value of a function of a random variable is defined as follows:

Discrete Random Variable:
\[ E[f(X)] = \sum_{x} f(x) P(X = x) \]

Continuous Random Variable:
\[ E[f(X)] = \int_{all \ x} f(x) P(X = x) \, dx \]

Properties of Expected Value:

- **Constants** - \( E(c) = c \) if \( c \) is constant
- **Indicators** - \( E(I_A) = P(A) \) where \( I_A \) is an indicator function
- **Constant Factors** - \( E(cX) = cE(X) \)
- **Addition** - \( E(X + Y) = E(X) + E(Y) \)
- **Multiplication** - \( E(XY) = E(X)E(Y) \) if \( X \) and \( Y \) are independent.
Another common property of random variables we are interested in is the Variance which measures the squared deviation from the mean.

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

One common simplification:

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

Standard Deviation:

\[ SD(X) = \sqrt{\text{Var}(X)} \]

Which gives us:

\[ \text{Var}(aX) = a^2 \text{Var}(X) \]
\[ \text{Var}(X + c) = \text{Var}(X) \]
\[ \text{Var}(c) = 0 \]

What about \( \text{Var}(X + Y) \)?

What is \( \text{Var}(aX + b) \) when \( a \) and \( b \) are constants?

This quantity is known as Covariance, and is roughly speaking a generalization of variance to two variables

\[ \text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[(X - \mu_x)(Y - \mu_y)] = E[XY + \mu_x\mu_y - X\mu_y - Y\mu_x] = E(XY) - \mu_x\mu_y \]
Properties of Covariance

- \( \text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y \)

- \( \text{Cov}(X, Y) = 0 \) if \( X \) and \( Y \) are independent

- \( \text{Cov}(X, c) = 0 \)

- \( \text{Cov}(X, X) = \text{Var}(X) \)

- \( \text{Cov}(aX, bY) = ab \text{Cov}(X, Y) \)

- \( \text{Cov}(X + a, Y + b) = \text{Cov}(X, Y) \)

Properties of Variance, cont.

A general formula for the variance of the linear combination of two random variables:

\[
\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + \text{Cov}(X, Y)
\]

\[
\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - \text{Cov}(X, Y)
\]

From which we can see that

\[
\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + \text{Cov}(X, Y)
\]

\[
\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - \text{Cov}(X, Y)
\]

Bernoulli Random Variable

For a completely general formula for the variances of a linear combination of \( n \) random variables:

\[
\text{Var} \left( \sum_{i=1}^{n} c_i X_i \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov}(c_i X_i, c_j X_j)
\]

\[
= \sum_{i=1}^{n} c_i^2 \text{Var}(X_i) + \sum_{i=1}^{n} \sum_{j=1 \atop i \neq j}^{n} c_i c_j \text{Cov}(X_i, X_j)
\]

Let \( X \sim \text{Bern}(p) \), what is \( E(X) \) and \( \text{Var}(X) \)?
Let $X \sim \text{Binom}(n, p)$, what is $E(X)$ and $\text{Var}(X)$?

We can redefine $X = \sum_{i=1}^{n} Y_i$ where $Y_1, \cdots, Y_n \sim \text{Bern}(p)$, and since we are sampling with replacement all $Y_i$ and $Y_j$ are independent.

### Hypergeometric Random Variable - $E(X)$ - cont.

Let $X \sim \text{Hypergeo}(N, m, n)$ then $X = B_1 + B_2 + \cdots + B_n$

$P(B_i) = m/N$ for all $i$
Let $X \sim \text{Hypergeo}(N, m, n)$, what is $E(X)$?

$$E(X) = \frac{n m}{N} \left(1 - \frac{m}{N}\right)$$

Let $X \sim \text{Hypergeo}(N, m, n)$, what is $\text{Var}(X)$?

$$\text{Var}(X) = var\left(\sum_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} \text{Var}(B_i) + \sum_{i=1}^{n} \sum_{j \neq i}^{n} \text{Cov}(B_i, B_j)$$

$$\sum_{i=1}^{n} \text{Var}(B_i) = \sum_{i=1}^{n} \frac{m(N-m)}{N^2} = \frac{nm(N-m)}{N^2}$$
Let $X \sim \text{Poisson}(\lambda)$, what is $E(X)$?

Let $X \sim \text{Poisson}(\lambda)$, what is $\text{Var}(X)$?

St. Petersburg Lottery

We start with $\$1$ on the table and a coin.

At each step: Toss the coin; if it shows Heads, take the money. If it shows Tails, I double the money on the table.

Let $X$ be the amount you win, what is $E(X)$?

Moments

Some definitions,

Raw moment:

$$\mu'_n = E(X^n)$$

Central moment:

$$\mu_n = E[(X - \mu)^n]$$

Normalized / Standardized moment:

$$\frac{\mu_n}{\sigma^n}$$
Moments

Common Moments of Interest

Zeroth Moment:

\[ \mu'_0 = \mu_0 = 1 \]

First Moment:

\[ \mu'_1 = E(X) = \mu \]
\[ \mu_1 = E(X - \mu) = 0 \]

Second Moment:

\[ \mu'_2 = E[(X - \mu)^2] = \text{Var}(X) \]
\[ \mu'_2 - (\mu'_1)^2 = \text{Var}(X) \]

Third Moment:

\[ \text{Skewness}(X) = \frac{\mu_3}{\sigma^3} \]

Fourth Moment:

\[ \text{Kurtosis}(X) = \frac{\mu_4}{\sigma^4} \]
\[ \text{Ex. Kurtosis}(X) = \frac{\mu_4}{\sigma^4} - 3 \]

Note that some moments do not exist, which is the case when \( E(X^n) \) does not converge.

Moment Generating Function

The moment generating function of a discrete random variable \( X \) is defined for all real values of \( t \) by

\[ M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} P(X = x) \]

This is called the moment generating function because we can obtain the moments of \( X \) by successively differentiating \( M_X(t) \) wrt \( t \) and then evaluating at \( t = 0 \).

\[ M_X(0) = E[e^0] = 1 = \mu'_0 \]
\[ M'_X(t) = \frac{d}{dt} E[e^{tx}] = E \left[ \frac{d}{dt} e^{tx} \right] = E[Xe^{tx}] \]
\[ M'_X(0) = E[Xe^0] = E[X] = \mu'_1 \]
\[ M''_X(t) = \frac{d^2}{dt^2} E[X^2e^{tx}] = E \left[ \frac{d^2}{dt^2} (Xe^{tx}) \right] = E[X^2 e^{tx}] \]
\[ M''_X(0) = E[X^2e^0] = E[X^2] = \mu'_2 \]

Third and Forth Central Moments

\[ \mu_3 = E[(X - \mu)^3] = E(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3) \]
\[ = E(X^3) - 3\mu E(X^2) + 3\mu^2 E(X) - \mu^3 \]
\[ = E(X^3) - 3\mu^2 \sigma^2 - \mu^3 \]

\[ \mu_4 = E[(X - \mu)^4] = E(X^4 - 4X^3\mu + 6X^2\mu^2 - 4X\mu^3 + \mu^4) \]
\[ = E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 4\mu^3 E(X) + \mu^4 \]
\[ = E(X^4) - 4\mu^2 (\sigma^2 + \mu^2) + 2\mu^4 + 4\mu^2 \sigma^2 + \mu^4 \]
\[ = \mu'_4 - 4\mu \mu'_3 + 6\mu^2 \sigma^2 + 3\mu^4 \]

Moment Generating Function - Poisson

Let \( X \sim \text{Pois} (\lambda) \) then
Moment Generating Function - Poisson Skewness

Moments

Moments

Moment Generating Function - Poisson Kurtosis

Moments

Moment Generating Function - Binomial

Let $X \sim \text{Binom}(n, p)$ then

\[
M_X(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\
= [pe^t + (1-p)]^n
\]

\[
M_X'(t) = npe^t (pe^t + 1 - p)^{n-1}
\]

\[
M_X'(0) = \mu_1' = np = E(X)
\]

\[
M_X''(t) = n(n-1)(pe^t)^2 (pe^t + 1-p)^{n-2} + npe^t (pe^t + 1-p)^{n-1} \\
M_X''(0) = \mu_2' = n(n-1)p^2 + np = E(X^2)
\]

\[
\text{Var}(X) = \mu_2' - (\mu_1')^2 = n(n-1)p^2 + np - n^2p^2 \\
= np((n-1)p + 1 - np) = np(1-p)
\]

Moments

Moment Generating Function - Normal

Let $Z \sim \mathcal{N}(0, 1)$ and $X \sim \mathcal{N}(\mu, \sigma^2)$ where $X = \mu + \sigma Z$ then

\[
M_X(t) = E[e^{tX}] = E[e^{t(\mu + \sigma Z)}] = E[e^{t\mu}e^{t\sigma Z}] \\
= e^{t\mu}E[e^{t\sigma Z}] = e^{t\mu}M_U(t\sigma) \\
= e^{t\mu}e^{t^2\sigma^2/2} = \exp \left( t\mu + \frac{t^2\sigma^2}{2} \right)
\]

Moments

Moments
Moment Generating Function - Normal, cont.

\[ M_X(t) = \exp(\mu t + \frac{t^2 \sigma^2}{2})(\mu + t\sigma^2) \]

\[ M_X(0) = \mu' = \mu \]

\[ M_X(t) = \exp \left( \mu t + \frac{t^2 \sigma^2}{2} \right) \sigma^2 + \exp \left( \mu t + \frac{t^2 \sigma^2}{2} \right) (\mu + t\sigma^2)^2 \]

\[ M_X(0) = \mu'' = \mu^2 + \sigma^2 \]

\[ \mu_3 = M_X'''(0) = \mu^3 + 3\mu \sigma^2 \]

\[ \mu_4 = M_X''''(0) = \mu^4 + 6\mu^2 \sigma^2 + 3\sigma^4 \]

Skewness \(X) = \frac{\mu_3}{\sigma^3} = \frac{\mu^3 - 3\mu \sigma^2 - \mu^3}{\sigma^3} = \frac{\mu^3 + 3\mu \sigma^2 - 3\mu \sigma^2 - \mu^3}{\sigma^3} = 0 \]

Kurtosis \(X) = \frac{\mu_4}{\sigma^4} = \frac{\mu_4 - 4\mu \mu_3 + 6\mu^2 \sigma^2 + 3\mu^4}{\sigma^4} = \frac{\mu_4 - 4\mu (\mu^3 + 3\mu \sigma^2) + 6\mu^2 \sigma^2 + 3\mu^4}{\sigma^4} = \frac{\mu^4 + 6\mu^2 \sigma^2 + 3\sigma^4 - 6\mu^2 \sigma^2 - 1\mu^4}{\sigma^4} = \frac{3\sigma^4}{\sigma^4} = 3 \]

Ex. Kurtosis \(X) = 0 \]