Moments

Definition. The **r$^{\text{th}}$ moment about the origin** of a random variable X is $\mu_r' = E(X^r)$.

Definition. The first moment about the origin of a random variable is called the **mean** and is denoted by μ.

Proposition. *If a and b are constants, then* $\boxed{E(aX+b) = aE(X)+b}$.

Definition. The **r$^{\text{th}}$ moment about the mean** of a random variable X is $\mu_r = E\left[(X-\mu)^r\right]$.

Definition. The second moment about the mean of a random variable is called the **variance** and is denoted by σ^2. The **standard deviation** of a random variable is $\sigma = \sqrt{\sigma^2}$.

Proposition (A calculating formula for the variance). $\boxed{\sigma^2 = \mu_2' - \mu^2 = E(X^2) - [E(X)]^2}$

Proposition. *If a and b are constants, then* $\boxed{V(aX+b) = a^2V(X)}$.

Proposition (A calculating formula for μ_3). $\boxed{\mu_3 = \mu_3' - 3\mu_2'\mu + 2\mu^3}$

Definition. The **moment-generating function** of a random variable X is $M_X(t) = E\left(e^{tX}\right)$.

Proposition. *If $M_X(t)$ is the moment-generating function of a random variable X, then* $\boxed{M_X^{(r)}(0) = \mu_r' = E(X^r)}$

Proposition. *If a and b are constants, then* $\boxed{M_{aX+b}(t) = e^{bt}M_X(at)}$.

Definition. If X and Y are jointly distributed random variables with means μ_X and μ_Y, respectively, then $E\left[(X-\mu_X)(Y-\mu_Y)\right]$ is called the **covariance** of X and Y and is denoted σ_{XY}, $\operatorname{cov}(X,Y)$, or $C(X,Y)$. If σ_X and σ_Y are the standard deviations of X and Y, respectively, then $\operatorname{cor}(X,Y) = \dfrac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}$ is the **correlation** of X and Y.

Proposition (A calculating formula for the covariance). $\boxed{cov(X,Y) = E(XY) - E(X)E(Y)}$

Proposition. *If X and Y are independent random variables, then $E(XY) = E(X)E(Y)$.*

Proposition. *If $X_1, X_2, \ldots, X_n$ are random variables and $a_1, a_2, \ldots, a_n$ are constants, then*

$$E\left(\sum_{i=0}^{n} a_iX_i\right) = \sum_{i=0}^{n} a_iE(X_i)$$

and

$$var\left(\sum_{i=0}^{n} a_iX_i\right) = \sum_{i=0}^{n} a_i^2 var(X_i) + 2\sum_{i<j}\sum a_ia_j\, cov(X_i, X_j)$$

Special discrete distributions

Definition. A random variable X has a **discrete uniform distribution** if it is equally likely to assume any one of a finite set of possible values.

Definition. A random variable X has a **Bernoulli distribution** with parameter θ (with $0 < \theta < 1$) if its probability distribution is

$$\boxed{f(x;\theta) = \begin{cases} 1-\theta & \text{if } x = 0 \\ \theta & \text{if } x = 1 \end{cases}}$$

The outcome 1 is often referred to as "success" while 0 is "failure" and the experiment is often called a Bernoulli trial.

Proposition. *The mean and variance of a Bernoulli random variable are $\mu = \theta$ and $\sigma^2 = \theta(1-\theta)$.*

Definition. The total number of successes in n independent, identically distributed (iid) Bernoulli trials is a random variable with a **Binomial distribution**. A random variable X has a binomial distribution with parameters n and θ if its probability distribution function is

$$\boxed{b(x;n,\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \text{ for } x = 0, 1, \ldots, n}$$

Proposition. *The mean and variance of a binomial distribution are $\mu = n\theta$ and $\sigma^2 = n\theta(1-\theta)$. The moment-generating function of a binomial distribution is $M_X(t) = [1 + \theta(e^t - 1)]^n$.*

Definition. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed (iid) Bernoulli trials, all with probability of success θ. Let N be the trial on which the first success occurs. The random variable N is said to have a **geometric distribution** with parameter θ and its probability distribution function is

$$\boxed{g(n;\theta) = \theta(1-\theta)^{n-1} \text{ for } n = 1, 2, 3, \ldots}$$

Proposition. *The mean and variance of a geometric distribution are $\mu = \frac{1}{\theta}$ and $\sigma^2 = \frac{1}{\theta}\left(\frac{1}{\theta} - 1\right)$.*

Definition. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed (iid) Bernoulli random variables, all with probability of success θ. Let N be the trial on which the k^{th} success occurs (so the possible values for N are $k, k+1, k+2, \ldots$). The random variable N is said to have a **negative binomial (or binomial waiting-time or Pascal) distribution** with parameters k and θ and its probability distribution function is

$$\boxed{b^*(n;k,\theta) = \binom{n-1}{k-1}\theta^k(1-\theta)^{n-k} \text{ for } n = k, k+1, k+2, \ldots}$$

Proposition. *The mean and variance of a negative binomial distribution are $\mu = \frac{k}{\theta}$ and $\sigma^2 = \frac{k}{\theta}\left(\frac{1}{\theta} - 1\right)$.*

Definition. A random variable with the probability distribution function

$$\boxed{p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \text{ for } x = 0, 1, 2, \ldots}$$

is said to have a **Poisson distribution** with parameter $\lambda > 0$.

Proposition. *The mean and variance of a Poisson distribution are $\mu = \lambda$ and $\sigma^2 = \lambda$. The moment-generating function of a Poisson random distribution is $M_X(t) = e^{\lambda(e^t - 1)}$.*

Definition. Suppose n elements are to be selected without replacement from a population of size N of which M are successes. The number of successes selected is a **hypergeometric** random variable and its probability distribution function is

$$\boxed{h(x;nN,M) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}}$$

Proposition. *The mean and variance of a hypergeometric distribution are $\mu = \frac{nM}{N}$ and $\sigma^2 = \frac{nM(N-M)(N-n)}{N^2(N-1)}$.*

Special continuous distributions

Definition. A random variable X has a **uniform continuous distribution** with parameters α and β (with $\alpha < \beta$) if and only if the following function is a probability density for X: $\boxed{f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{elsewhere} \end{cases}}$.

Definition. The **gamma function** is defined as $\boxed{\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}dy}$ for $\alpha > 0$.

Proposition. *For any positive integer n, $\Gamma(n) = (n-1)!$*

Definition. A random variable X has a **gamma distribution** with parameters $\alpha > 0$ and $\beta > 0$ if and only if the following function is a probability density for X: $\boxed{g(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}}$.

Proposition. *A gamma distribution with parameters α and β has moment-generating function $M_X(t) = (1-\beta t)^{-\alpha}$, mean $\mu = \alpha\beta$, and variance $\sigma^2 = \alpha\beta^2$.*

Definition. A random variable X has an **exponential distribution** with parameter $\theta > 0$ if and only if the following function is a probability density for X: $\boxed{g(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & \text{if } x > 0 \\ 0 & \text{elsewhere} \end{cases}}$ (a gamma distribution with $\alpha = 1$ and $\beta = \theta$).

Proposition. *An exponential distribution with parameter θ has mean $\mu = \theta$ and variance $\sigma^2 = \theta^2$.*

Definition. A random variable X has a **chi-square distribution** with parameter $\nu > 0$ if and only if the following function is a probability density for X: $\boxed{f(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\frac{\nu-2}{2}}e^{-\frac{\nu}{2}} & \text{if } x > 0 \\ 0 & \text{elsewhere} \end{cases}}$ (a gamma distribution with $\alpha = \frac{\nu}{2}$ and $\beta = 2$).

Proposition. *A chi-square distribution with parameter ν has mean $\mu = \nu$ and variance $\sigma^2 = 2\nu$.*

Definition. A random variable X has a **beta distribution** with parameters $\alpha > 0$ and $\beta > 0$ if and only if the following function is a probability density for X: $\boxed{f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}}$.

Proposition. *A beta distribution with parameters α and β has mean $\mu = \dfrac{\alpha}{\alpha+\beta}$ and variance $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.*

Definition. A random variable X has a **normal distribution** with parameters μ and $\sigma > 0$ if and only if the following function is a probability density for the X: $\boxed{n(x;\mu,\sigma) = \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \text{ for all } x \in \mathbb{R}}$.

Proposition. *A normal distribution with parameters μ and σ has a moment-generating function $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$, mean $\mu = \mu$ and variance $\sigma^2 = \sigma^2$.*