Probability Reference

Combinatorics and Sampling

- A **permutation** is an *ordered* selection. The number of permutations of k items picked from a list of n items, *without replacement*, is

$$P(n,k) := \underbrace{n(n-1)(n-2)\cdots(n-k+1)}_{k\text{-factors}} = \frac{n!}{(n-k)!} =: (n)_k$$

  When selecting *with replacement*, the number of possibilities is n^k.

- A **combination** is an *unordered* selection. The number of combinations of k items chosen from a list of n different items, *without replacement*, is

$$C(n,k) := \frac{P(n,k)}{k!} =: \binom{n}{k} = \underbrace{\frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)(k-2)\cdots 1}}_{k\text{-factors, numerator and denominator}} = \frac{n!}{k!(n-k)!}$$

  The number of ways to select k objects from n different items *with replacement* is

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle := \binom{n+k-1}{k} = \underbrace{\frac{n(n+1)(n+2)\cdots(n+k-1)}{k(k-1)(k-2)\cdots 1}}_{k\text{-factors, above and below}}$$

  (This is also the number of nonnegative integer solutions of the equation $x_1 + x_2 + \cdots + x_n = k$ and the number of ways to distribute k identical objects into n distinct boxes.)

- The number of distinct ways of distributing n objects into k distinct classes of size $n_1, n_2, \ldots, n_k$, *without replacement* and with no order *within* each class, is

$$\binom{n}{n_1, n_2, \ldots, n_k} := \frac{n!}{n_1! n_2! \cdots n_k!}, \text{ where } n_1 + n_2 + \cdots + n_k = n$$

- **Binomial Theorem**:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

- **Multinomial Theorem**:

$$(a_1 + a_2 + \cdots + a_k)^n = \sum \binom{n}{n_1, n_2, \ldots, n_k} a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k},$$

  where the sum is taken over all nonnegative integer values of $n_1, n_2, \ldots, n_k$ for which $n_1 + n_2 + \cdots + n_k = n$.

- **Stirling's Formula**: $n! \doteq \sqrt{2\pi n}\,(n/e)^n$ or more accurately

$$n! \doteq \sqrt{2\pi} \left(n + \frac{1}{2}\right)^{n+(1/2)} e^{-n}$$

- Binomial coefficient identities:

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}, \quad \binom{-n}{k} = (-1)^k \binom{n+k-1}{k}, \text{ for } n > 0, \quad \binom{n}{k}\binom{n-k}{p} = \binom{n}{p}\binom{n-p}{k} = \binom{n}{k-p}\binom{n-k+p}{p},$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \quad \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0, \quad \sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1},$$

$$\sum_{k=0}^{r} \binom{n+k}{k} = \binom{n+r+1}{r}, \quad \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r}, \quad \sum_{k=0}^{m} \binom{m}{k}\binom{n}{r+k} = \binom{m+n}{m+r}$$

- Some useful series:

$\sum_{k=1}^{n} k = \frac{1}{2}n\,(n+1)$ $\qquad \sum_{k=1}^{n} k^2 = \frac{1}{6}n\,(n+1)\,(2n+1)$ $\qquad \sum_{k=1}^{n} k^3 = \frac{1}{4}n^2\,(n+1)^2$

$\sum_{k=m}^{n} r^k = \dfrac{r^m - r^{n+1}}{1-r}$ $\qquad \sum_{k=0}^{\infty} r^k = \dfrac{1}{1-r}$, for $|r|<1$ $\qquad \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$

$\sum_{k=1}^{\infty} \frac{x^k}{k} = -\log\,(1-x)$, for $|x|<1$ $\qquad \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = \cosh x$ $\qquad \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = \sinh x$

$(1-t)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} t^k = \sum_{k=0}^{\infty} \binom{n+k-1}{n-1} t^k$, $|t|<1$

Probability

If $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are **events** (defined as subsets of the **sample space** $\mathcal{S}$ of all possible outcomes of an experiment) then Pr is a **probability measure**, when the following are true:

(i) $0 \le \Pr(\mathcal{A}) \le 1$, (ii) $\Pr\left(\bigcup_1^\infty \mathcal{A}_i\right) = \sum_{i=1}^{\infty} \Pr\left(\mathcal{A}_i\right)$, for pairwise disjoint $\mathcal{A}_i$; (iii) $\Pr(\mathcal{S}) = 1 \Leftrightarrow \Pr\left(\varnothing\right) = 0$.

- The complement of an event is defined to be $\mathcal{A}' = \{x : x \notin \mathcal{A}\}$, then $\Pr\left(\mathcal{A}'\right) = 1 - \Pr\left(\mathcal{A}\right)$ is the **Law of Complements**; $\Pr\left(\mathcal{A}\cup\mathcal{B}\right) = \Pr\left(\mathcal{A}\right) + \Pr\left(\mathcal{B}\right) - \Pr\left(\mathcal{A}\cap\mathcal{B}\right)$ is the **Principle of Inclusion-Exclusion**.
- **Conditional probability** of $\mathcal{A}$ given $\mathcal{B}$,

$$\Pr\left(\mathcal{A}|\mathcal{B}\right) := \frac{\Pr\left(\mathcal{A}\cap\mathcal{B}\right)}{\Pr\left(\mathcal{B}\right)}, \text{ when } \Pr\left(\mathcal{B}\right) > 0;$$

  this implies $\Pr\left(\mathcal{A}\cap\mathcal{B}\right) = \Pr\left(\mathcal{A}|\mathcal{B}\right)\Pr\left(\mathcal{B}\right) = \Pr\left(\mathcal{B}|\mathcal{A}\right)\Pr\left(\mathcal{A}\right) = \Pr\left(\mathcal{B}\cap\mathcal{A}\right)$.

- The events $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are **independent** if

$$\Pr\left(\mathcal{A}_{r_1}\cap\mathcal{A}_{r_2}\cap\cdots\cap\mathcal{A}_{r_k}\right) = \Pr\left(\mathcal{A}_{r_1}\right)\Pr\left(\mathcal{A}_{r_2}\right)\cdots\Pr\left(\mathcal{A}_{r_k}\right),$$

  for $\{r_1, r_2, \ldots, r_k\}$ any subset of $1:n$. This implies that $\mathcal{A}^\#_{i_1}, \mathcal{A}^\#_{i_2}, \ldots, \mathcal{A}^\#_{i_s}$ are independent, where $\mathcal{A}^\#$ can be either $\mathcal{A}$ or $\mathcal{A}'$, separately for each set as $k = 2:n$.

- If $\mathfrak{B} = \{\mathcal{B}_k, k = 1:n\}$ is a **partition** of the sample space $\mathcal{S}$, meaning $\mathcal{B}_i \cap \mathcal{B}_j = \varnothing$ for $i \ne j$ and $\bigcup_{i=1}^n \mathcal{B}_i = \mathcal{S}$, then the **Law of Total Probability** says

$$\Pr\left(\mathcal{A}\right) = \sum_{i=1}^{n} \Pr\left(\mathcal{A}|\mathcal{B}_i\right)\Pr\left(\mathcal{B}_i\right),$$

  and **Bayes' formula** is

$$\Pr\left(\mathcal{B}_k|\mathcal{A}\right) = \frac{\Pr\left(\mathcal{A}|\mathcal{B}_k\right)\Pr\left(\mathcal{B}_k\right)}{\sum_{i=1}^{n}\Pr\left(\mathcal{A}|\mathcal{B}_i\right)\Pr\left(\mathcal{B}_i\right)}$$

Discrete Random Variables

1. X has **probability mass function** pmf $f(x)$ if (i) $f(x) \ge 0$, (ii) $\sum f(x) = 1$, (iii) $f(x_k) = \Pr(X = x_k)$.
2. X has **cumulative distrbution function** cdf $F(x)$ if $F(x) := \Pr(X \le x) = \sum_{y \le x} f(y)$; $\Pr(a < x \le b) = F(b) - F(a)$; $f(x_k) = F(x_k) - F(x_{k-1})$.

Continuous Random Variables

1. X has **probability density function** pdf $f(x)$ if (i) $f(x) \ge 0$, (ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$, (iii) $\Pr(a < x \le b) = \int_a^b f(x)\,dx$.
2. X has cdf $F(x)$ if $F(x) := \int_{-\infty}^{x} f(\xi)\,d\xi$; $\Pr(a \le X < b) = F(b) - F(a)$; $f(x) = \dfrac{dF}{dx}$.
3. The **median** $\tilde{x}$ satisfies $F\left(\tilde{x}\right) = \frac{1}{2}$ and the p^{th} **percentile** x_p satisfies $F\left(x_p\right) = p$. The **interquartile range** is $IQR := x_{0.75} - x_{0.25}$ and the **interdecile range** is $IDR := x_{0.90} - x_{0.10}$

Discrete and Continuous cdfs

$F(x)$ is (i) nondecreasing, (ii) $\lim_{x\to-\infty} F(x)=0$, (iii) $\lim_{x\to\infty} F(x)=1$, and (iv) $F(x)$ is right continuous.

Independent and Exchangeable Random Variables

The rvs $X_1, X_2, \ldots, X_n$ are **independent** if and only if the joint pf is the product of the marginals, i.e.,

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$

The rvs $X_1, X_2, \ldots, X_n$ are **exchangeable** if and only if the joint pf is invariant under interchanges of its arguments, i.e.,

$$f(x_1, x_2, \ldots, x_n) = f(x_{i_1}, x_{i_2}, \ldots, x_{i_n}),$$

for any permutation $(i_1, i_2, \ldots, i_n)$ of $1:n$. Independent and identically distributed (iid) rvs are exchangeable, but independence and exchangeability, although overlapping in some areas, are distinct concepts. For instance, independence *does not* imply exchangeability. *Nor* does exchangeability imply independence.

Expectation Values

By "definition," the **expectation** of a function of a rv is

$$E(g(X)) := \begin{cases} \sum\limits_{\text{range}(X)} g(x)f(x), & X \text{ discrete,} \\ \int_{-\infty}^{\infty} g(x)f(x)\,dx, & X \text{ continuous.} \end{cases}$$

For rvs of the mixed type with **probability function** pf $f(x) = \alpha f_{\text{disc}}(x) + (1-\alpha) f_{\text{cont}}(x)$, you can *only* define the moments

$$E\left(X^k\right) = \alpha \sum_{\text{discrete}(X)} x^k f_{\text{disc}}(x) + (1-\alpha) \int_{\text{continuous}(X)} x^k f_{\text{cont}}(x)\,dx$$

1. r^{th} **moment** is $\mu'_r := E(X^r)$; the **mean** is $\mu := E(X)$; r^{th} **central moment** is $\mu_r := E(X-\mu)^r$; r^{th} **absolute deviation** is $\nu_r := E(|X-\mu|^r)$;

   the **variance** is $\text{var}(X) := \sigma^2 := \mu_2 = E(X-\mu)^2 = E\left(X^2\right) - \mu^2$.

$$\mu_r = \mu'_r - \binom{r}{1}\mu\mu'_{r-1} + \binom{r}{2}\mu^2\mu'_{r-2} + \cdots + (-1)^{r-1}(r-1)\mu^r$$

$$\mu'_r = \mu_r + \binom{r}{1}\mu\mu_{r-1} + \binom{r}{2}\mu^2\mu_{r-2} + \cdots + \binom{r}{r-2}\mu^{r-2}\mu_2 + \mu^r$$

2. **Coefficient of skewness** is $\gamma_1 := E((X-\mu)/\sigma)^3$ and **coefficient of excess** is $\gamma_2 = E((X-\mu)/\sigma)^4 - 3$.

3. $E\left(\sum a_k X_k\right) = \sum a_k E(X_k)$; $\text{var}\left(\sum a_k X_k\right) = \sum a_k^2 \text{var}(X_k) + 2\sum\sum_{j<k} a_j a_k \text{cov}(X_j, X_k)$. The **covariance** and **correlation** are defined by:

$$\text{cov}(X_j, X_k) := E\big((X_j-\mu_j)(X_k-\mu_k)\big) = E(X_j X_k) - \mu_j\mu_k =: \sigma_{jk}; \quad \rho_{jk} := \text{corr}(X_j, X_k) = \frac{\sigma_{jk}}{\sigma_j\sigma_k}$$

$$\text{cov}\left(\sum a_i X_i, \sum b_j X_j\right) = \sum a_i b_i \,\text{var}(X_i) + \sum_{i<j}\sum (a_i b_j + a_j b_i)\,\text{cov}(X_i, X_j)$$

4. **Conditional expectations**: $E(Y) = E(E(Y|X))$ and

$$\text{var}(Y) = E(\text{var}(Y|X)) + \text{var}(E(Y|X)),$$

   where $\text{var}(Y|X) := E\left((Y - E(Y|X))^2 | X\right)$.

5. When X and Y are independent rvs,

$$\text{var}(XY) = \text{var}(X)\,\text{var}(Y) + E^2(X)\,\text{var}(Y) + E^2(Y)\,\text{var}(X)$$

Generating Functions

- **Moment generating function**, mgf: $M_X(t) := E\left(e^{tX}\right)$, $M_X^{(n)}(0) = \mu'_n$, $M_{aX+b}(t) = e^{bt}M_X(at)$, $\mu = M'_X(0)$, $\sigma^2 = M''_X(0) - \left(M'_X(0)\right)^2$.
- **Cumulant generating function**, cgf: $K_X(t) := \log\left(M_X(t)\right)$, $K'_X(0) = \mu$, $K''_X(0) = \sigma^2$, $K'''_X(0) = E\left(X-\mu\right)^3$, the r^{th} cumulant is $\kappa_r = K_X^{(r)}(0) = E\left(X-\mu\right)^r$.
- **Factorial generating function**, fgf: $P_X(s) := E\left(s^X\right)$, $P_X^{(r)}(1) = \mu_{[r]} := E\left(X(X-1)\cdots(X-r+1)\right)$.
- **Characteristic function**, cf: $\varphi_X(\omega) := E\left(e^{i\omega X}\right)$, $\varphi_X^{(n)}(0) = i^n\mu'_n$, $\varphi_{aX+b}(\omega) = e_X^{ibt}\varphi(a\omega)$.

Order Statistics

A **random sample** of size n is a set $\{X_1, X_2, \ldots, X_n\}$ of **independent and identically distributed** (iid) rvs. The **order statistics** of the random sample are defined to be $X_{(1;n)} \leq X_{(2;n)} \leq \cdots \leq X_{(n;n)}$. We assume they are drawn from a population with pdf $f(x)$ and cdf $F(x)$.

1. The pdf and cdf of $Y = X_{(r;n)}$ are given by

$$g_r(y) = \binom{n}{r-1, 1, n-r} \left[F(y)\right]^{r-1} f(y) \left[1 - F(y)\right]^{n-r},$$

$$G_r(y) = \sum_{i=r}^{n} \binom{n}{i} \left[F(y)\right]^{i} \left[1 - F(y)\right]^{n-i}$$

The joint pdfs of two order statistics, $Y_r = X_{(r;n)} \leq Y_s = X_{(s;n)}$ are given by

$$g_{r,s}(y_r, y_s) = \binom{n}{r-1, 1, s-r-1, 1, n-s} \left[F(y_r)\right]^{r-1} f(y_r) \left[F(y_s) - F(y_r)\right]^{s-r-1} f(y_s) \left[1 - F(y_s)\right]^{n-s}, \; y_r \leq y_s$$

and the pdf of all the order statistics is

$$g(y_1, y_2, \ldots, y_n) = n!\, f(y_1, y_2, \ldots, y_n) \text{ for } y_1 \leq y_2 \leq \cdots \leq y_n$$

2. The pdf and cdf of the **range**, $R := X_{(n;n)} - X_{(1;n)}$, are given by

$$f_R(r) = n(n-1)\int_{-\infty}^{\infty} \left[F(x+r) - F(x)\right]^{n-2} f(x) f(x+r)\, dx,$$

$$F_R(r) = n\int_{-\infty}^{\infty} \left[F(x+r) - F(x)\right]^{n-1} f(x)\, dx$$

Transformation of Variables

- If $Y = u(X)$ is a smooth one-to-one transformation, then

$$G(y) = \Pr\left(Y \leq y\right) = \Pr\left(u(X) \leq y\right) = \Pr\left(X \leq u^{-1}(y)\right) = F\left(u^{-1}(y)\right)$$

The corresponding pdf is the derivative: $g(y) = f\left(x(y)\right)\left|\frac{dx}{dy}\right|$. If the transformation is not one-to-one, break up its support into a union of intervals over each of which it is one-to-one and apply the previous formula to each piece and sum the result. E.g.,

$$Y = X^2, G(y) = \Pr\left(Y \leq y\right) = \Pr\left(X^2 \leq y\right) = \Pr\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) = F\left(\sqrt{y}\right) - F\left(-\sqrt{y}\right),$$

so that

$$g(y) = \frac{1}{2\sqrt{y}}\left(f\left(\sqrt{y}\right) + f\left(-\sqrt{y}\right)\right)$$

For discrete rvs, the pmf is $g(y_k) = F\left(u^{-1}(y_k)\right) - F\left(u^{-1}(y_{k-1})\right) = f\left(u^{-1}(y_k)\right)$.

You should know that for *any* continuous rv, both $U = F(X)$ and $V = 1 - F(X)$ are $\mathcal{U}nif\,(0,1)$.

- If $\mathbf{Y} := [Y_1, Y_2, \ldots, Y_n] = [u_1(X_1, X_2, \ldots, X_n), \ldots, u_n(X_1, X_2, \ldots, X_n)]$ is a smooth invertible multivariate transformation, then use the **Jacobian Change of Variable Theorem** to write,

$$g(y_1, y_2, \ldots, y_n) = f(x_1(\mathbf{y}), x_2(\mathbf{y}), \ldots, x_n(\mathbf{y})) \left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)} \right|,$$

  where the **Jacobian** is defined to be

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)} := \det \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{bmatrix}$$

- For sums of independent rvs, use the mgf result: If $S = X_1 + X_2 + \cdots + X_n$, then

$$M_S(t) = M_{X_1}(t)M_{X_2}(t) \cdots M_{X_n}(t),$$

  which for the iid case reduces to $M_S(t) = M_X^n(t)$.

- Also for sums of random variables, the pdf, $f(s)$, of the sum is related to the pdfs of the individual X_i, $p_i(x)$, via the convolution product $f = p_1 * p_2 * \cdots * p_n$, where the product is defined recursively by

$$(p_1 * p_2)(x) := \int_{-\infty}^{\infty} p_1(x - y)p_2(y)\, dy,$$

  $p_1 * p_2 = p_2 * p_1$, and $p_1 * (p_2 * p_3) = (p_1 * p_2) * p_3$. This is usually not very useful except for distributions for which $f(s)$ can be more easily calculated other ways, e.g., mgfs. (See the last section of this reference sheet.)

Definitions and Results

- If X_n has cdf $F_n(x)$ for $n = 1 : \infty$ and if for some cdf $F(x)$ we have $\lim_{n\to\infty} F_n(x) = F(x)$ for all values of x at which $F(x)$ is continuous, then the sequence $\{X_n\}$ **converges in distribution** to X, which has cdf $F(x)$, and we write $X_n \overset{d}{\to} X$.

- If X_n has mgf $M_n(t)$, X has mgf $M(t)$, and there is an $a > 0$ such that $\lim_{n\to\infty} M_n(t) = M(t)$ for all $t \in (-a, a)$, then $X_n \overset{d}{\to} X$.

- We say the sequence $\{X_n\}$ **converges stochastically** to a constant c if the limiting distribution puts all its mass at the atom $\{c\}$, written $X_n \overset{P}{\to} c$.

- The sequence $\{X_n\}$ **converges in probability** to X if $\lim_{n\to\infty} \Pr(|X_n - X| < \varepsilon) = 1$, for any $\varepsilon > 0$. This is written as $X_n \overset{P}{\to} X$.

- If $\Omega_0 := \{\omega : \lim_{n\to\infty} X_n(\omega) = X \text{ exists}\}$ and $\Pr(\Omega_0) = 1$, then we say that X_n **converges almost surely** and we write $X_n \overset{\text{a.s.}}{\to} X$.

- Slutsky's Theorem says: (a) If $X_n \overset{P}{\to} X$, then $X_n \overset{d}{\to} X$. (b) If $X_n \overset{P}{\to} c$, then $g(X_n) \overset{P}{\to} g(c)$, whenever g is continuous at c. (c) If $X_n \overset{d}{\to} X$ and $Y_n \overset{P}{\to} c$, then (i) $X_n + Y_n \overset{d}{\to} X + c$, (ii) $X_nY_n \overset{d}{\to} Xc$, (iii) $X_n/Y_n \overset{d}{\to} X/c$. (d) If $X_n \overset{d}{\to} X$, then for any continuous function $g(y)$, $g(X_n) \overset{d}{\to} g(X)$.

- **Central Limit Theorem**: (Form 1) If $X_1, X_2, \ldots, X_n$ are iid from a distribution with mean μ and variance $\sigma^2 < \infty$, then

$$\lim_{n\to\infty} Z_n := \lim_{n\to\infty} \frac{\sum_1^n X_i - n\mu}{\sigma\sqrt{n}} = Z \sim \mathcal{N}(0, 1)$$

  (Form 2) If as above, then

$$\lim_{n\to\infty} Z_n := \lim_{n\to\infty} \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = Z \sim \mathcal{N}(0, 1)$$

(**Berry-Esseen Bound**) If, in addition, $(E|X_i|)^{2+\delta} = \gamma^{2+\delta} < \infty$, for some $\delta \in (0,1]$, then there is a constant c_δ such that

$$\sup\left\{\left|\Pr\left(\bar{X} - \mu < z\frac{\sigma}{\sqrt{n}}\right) - \Phi(z)\right| : x \in \mathbb{R}\right\} \le \frac{c_\delta}{n^{\delta/2}}\left(\frac{\gamma}{\sigma}\right)^{2+\delta}$$

The $\delta = 1$ case is most often cited: $(E|X_i|)^3 = \gamma^3$ yields

$$\sup\left\{\left|\Pr\left(\bar{X} - \mu < z\frac{\sigma}{\sqrt{n}}\right) - \Phi(z)\right| : x \in \mathbb{R}\right\} \le \frac{c_1}{\sqrt{n}}\left(\frac{\gamma}{\sigma}\right)^3,$$

and $c_1 \le 1.322$.

Special Discrete Random Variables

1. $\mathcal{B}in(n,p)$, **Binomial**: $X = \#$ successes in n, a fixed number of independent Bernoulli trials with constant $p = \Pr(\text{Success}) =: 1 - q$.

$$b(x;n,p) = \binom{n}{x} p^x q^{n-x},\ x = 0:n;$$

$$\mu = np;\ \sigma^2 = npq;\ \mu_{(r)} = (n)_r\, p^r;\quad M_X(t) = (pe^t + q)^n.$$

2. $\mathcal{H}yper(n,N,k)$, **Hypergeometric**: $X = \#$ defectives in sampling n items without replacement from a set of N items of which D are defectives.

$$h(x;n,N,D) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}},\quad x = \max\{0, n+D-N\} : \min\{D,n\}$$

$$\mu = n\left(\frac{D}{N}\right);\quad \sigma^2 = \left(\frac{N-n}{N-1}\right) n \left(\frac{D}{N}\right)\left(1 - \frac{D}{N}\right);\quad \mu_{[r]} = \frac{n^{[r]}D^{[r]}}{N^{[r]}}$$

3. $\mathcal{P}ois(\theta)$, **Poisson**: $X = \#$ of occurrences of events occurring "randomly and independently" in a time T and at a rate λ when $\theta = \lambda T$.

$$p(x;\theta) = \frac{\theta^x}{x!}e^{-\theta},\ x = 0:\infty;\quad \mu = \sigma^2 = \theta;\quad \mu_2' = \theta(1+\theta)$$

$$\mu_{[r]} = \theta^r;\quad M_X(t) = \exp\left\{\theta\left(e^t - 1\right)\right\}$$

The **Law of Rare Events** tells us that the limit of $\mathcal{B}in(n,p)$ as $n \to \infty$, $p \to 0$, and $np = \theta$ is $\mathcal{P}ois(\theta)$.

4. $\mathcal{B}in^*(r,p)$, **Negative Binomial**: $X = \#$ of trials until r^{th} success, or $N\mathcal{B}in(r,p)$: $Y = \#$ failures until r^{th} success$= X - r$.

$$b^*(x;r,p) = \binom{x-1}{r-1} p^r q^{x-r},\ x = r:\infty;\quad \mu_X = \frac{r}{p};\quad \sigma_X^2 = \frac{rq}{p^2};$$

$$f(y) = \binom{-r}{y} p^r(-q)^y = \binom{r+y-1}{y} p^r q^y,\ y = 0:\infty;\quad \mu_Y = \frac{rq}{p};\quad \sigma_Y^2 = \frac{rq}{p^2}$$

$$M_X(t) = \frac{p^r}{(e^{-t} - q)^r};\quad M_Y(t) = \frac{p^r}{(1 - qe^t)^r}$$

   (a) $\mathcal{G}eo(p)$, **Geometric**: $X = \#$ of trials until the first success. This is $\mathcal{B}in^*(1,p)$. So, $f(x;p) = q^{x-1}p$, for $x = 1:\infty$. $\mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$, $M_X(t) = p\left(e^{-t} - q\right)^{-1}$.

5. $\mathcal{M}ult(\mathbf{n},\mathbf{p})$, **Multivariate**: $X_i = \#$ of occurences falling into category i when the probability of having an outcome in each category is the same for each independent trial.

$$\Pr(\mathbf{X} = \mathbf{x}) = \binom{n}{\mathbf{x}}\mathbf{p}^{\mathbf{x}} := \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},\ \sum_1^k x_i = n \text{ and } \sum_1^k p_i = 1,$$

and $E(X_i) = np_i$, $\operatorname{var}(X_i) = np_i(1 - p_i)$, and $\operatorname{cov}(X_i, X_j) = -np_ip_j$ for $i \ne j$.

Special Continuous Random Variables

The **indicator function** is defined by

$$I_{(a,b)}(x) = \begin{cases} 1, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$$

1. $\mathcal{U}nif(a,b)$, **Uniform**: $f(x;a,b) = \dfrac{1}{b-a} I_{(a,b)}(x)$; $\mu = \tilde{x} = \frac{1}{2}(a+b)$; $\sigma^2 = \frac{1}{12}(b-a)^2$;

$$M_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)};\ \mu_r' = \frac{b^{r+1} - a^{r+1}}{(b-a)(r+1)};\ \gamma_1 = 0;\ \gamma_2 = -\frac{6}{5}$$

2. $\mathcal{N}(\alpha, \beta^2)$, **Gaussian** or **Normal**: $n(x;\alpha,\beta^2) = \dfrac{1}{\beta}\phi\left(\dfrac{x-\alpha}{\beta}\right) := \dfrac{1}{\sqrt{2\pi\beta^2}} \exp\left\{-\dfrac{1}{2}\left(\dfrac{x-\alpha}{\beta}\right)^2\right\}$,

$$\mu = \tilde{x} = \alpha; \quad \sigma^2 = \beta^2; \quad \gamma_1 = 0;\ \gamma_2 = 0;\ M_X(t) = \exp\left(\alpha t + \frac{1}{2}\beta^2 t^2\right)$$

For the **standard normal**, $Z \sim \mathcal{N}(0,1)$, the pdf is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

3. $\log\mathcal{N}(\alpha,\beta)$, **log Normal**: $\mu = \exp\left(\alpha + \frac{1}{2}\beta^2\right)$; if $\omega = e^{\beta^2}$, then $\sigma^2 = \omega(\omega-1)e^{2\alpha}$; $\gamma_1 = (\omega+2)\sqrt{\omega-1}$, $\gamma_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 6$, $\tilde{x} = e^{\alpha}$; $\mu_r' = \exp\left(r\alpha + \frac{1}{2}r^2\beta^2\right)$

$$f(x;\alpha,\beta) = \frac{1}{x\sqrt{2\pi\beta^2}} \exp\left\{-\frac{1}{2\beta^2}(\log x - \alpha)^2\right\} I_{(0,\infty)}(x)$$

4. $inv\mathcal{G}(\alpha,\beta)$, **inverse Gaussian** or **inverse Normal**: $\mu = \alpha$, $\sigma^2 = \alpha^3\beta$, $\gamma_1 = 3\sqrt{\alpha\beta}$, $\gamma_2 = 15\alpha\beta$,

$$f(x;\alpha,\beta) = \frac{1}{\sqrt{2\pi\beta x^3}} \exp\left\{-\frac{(x-\alpha)^2}{2\alpha^2\beta x}\right\} I_{(0,\infty)} \text{ when } \alpha > 0 \text{ and } \beta > 0$$

and the mgf is

$$M_X(t) = \exp\left\{\frac{1}{\alpha\beta}\left[1 - \sqrt{\beta(1+\alpha^2 t)}\right]\right\};\ \kappa_r = (2r-3)!!\alpha^{2r-1}\beta^{r-1}$$

where the **semifactorial** is defined by

$$\begin{aligned} (2r)!! &= 2\cdot 4\cdot 6\cdot\cdots\cdot(2r) \\ (2r-1)!! &= 1\cdot 3\cdot 5\cdot\cdots\cdot(2r-1) \end{aligned}$$

5. $\mathcal{C}auchy(\alpha,\beta)$, **Cauchy**: $f(x) = \dfrac{\beta}{\pi}\dfrac{1}{\beta^2 + (x-\alpha)^2}$; $\beta > 0$, μ and σ^2 do not exist but $\tilde{x} = \alpha$ and the characteristic function

$$\varphi_X(\omega) = \exp\left(i\alpha\omega - \frac{|t|}{\beta}\right)$$

is the only generating function that exists. The parameter β is one half the interquartile range, i.e., $\beta = \frac{1}{2}IQR = \frac{1}{2}(Q_3 - Q_1) = \frac{1}{2}(x_{0.75} - x_{0.25})$.

6. $\mathcal{E}xp(\beta) = \mathcal{G}am(1,\beta)$, **Exponential**: $f(x) = \beta^{-1}e^{-x/\beta} I_{(0,\infty)}(x)$, $\beta > 0$, $\mu = \beta$; $\sigma^2 = \beta^2$; $\mu_r' = \beta^r r!$;

$$M_X(t) = (1-\beta t)^{-1}.$$

This is the distribution of the time until the next occurrence of a random event.

7. $\mathcal{G}am(\alpha,\beta)$, **Gamma**: $f(x;\alpha,\beta) = \dfrac{1}{\beta^\alpha\Gamma(\alpha)}\, x^{\alpha-1}e^{-x/\beta}\, I_{(0,\infty)}(x)$, $\alpha,\beta>0$, $\mu=\alpha\beta$; $\sigma^2=\alpha\beta^2$; $\gamma_1 = 2/\sqrt{\alpha}$, $\gamma_2 = 6/\alpha$,

$$\mu_r' = \frac{\Gamma(r+\alpha)}{\Gamma(\alpha)}\,\beta^r; \quad M_X(t) = (1-\beta t)^{-\alpha}$$

If α is a positive integer, then this is the distribution of the time until the α^{th} occurrence of a random event.

8. $\chi_n^2 = \mathcal{G}am(n/2,2)$, **Chi-Square**: $f(x;n) = \dfrac{1}{2^{n/2}\Gamma(n/2)}\, x^{(n/2)-1}e^{-x/2}\, I_{(0,\infty)}(x)$, for $n>0$; $\mu=n$;

$$\sigma^2 = 2n;\ \text{Mode} = n-2;\ \ \mu_r' = \frac{2^r\Gamma\left(\frac{n}{2}+r\right)}{\Gamma\left(\frac{n}{2}\right)};\ \ \gamma_1 = \sqrt{\frac{8}{n}},\ \gamma_2 = \frac{12}{n};\ M_X(t) = (1-2t)^{-n/2}$$

9. $\mathcal{B}eta(\alpha,\beta)$, **Beta**: $f(x;\alpha,\beta) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}\, I_{(0,1)}(x)$ for $\alpha,\beta>0$; $\mu = \dfrac{\alpha}{\alpha+\beta}$;

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2};\ \ \mu_r' = \frac{B(\alpha+r,\beta)}{B(\alpha,\beta)} = \prod_{k=0}^{r-1}\left(\frac{\alpha+k}{\alpha+\beta+k}\right)$$

10. $\mathcal{W}eib(\alpha,\beta)$, **Weibull**: $f(x;\alpha,\beta) = \alpha\beta x^{\beta-1}\exp\left(-\alpha x^\beta\right)\, I_{(0,\infty)}(x)$; $\mu = \alpha^{-1/\beta}\Gamma\left(1+\frac{1}{\beta}\right)$;

$$\mu_r' = \alpha^{-r/\beta}\Gamma\left(1+\frac{r}{\beta}\right);\ \ \sigma^2 = \alpha^{-2/\beta}\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right\}$$

11. $\mathcal{L}ap(\alpha,\beta)$, **Laplace** or **Double Exponential**: $f(x;\alpha,\beta) = \dfrac{1}{2\beta}\exp\{-|x-\alpha|/\beta\}$; $\mu=\tilde{x}=\alpha$;

$$\sigma^2 = 2\beta^2;\ \ \gamma_1 = 0;\ \gamma_2 = 3;\ \mu_{2r} = (2r)!\,\beta^r;\ \ M_X(t) = \frac{e^{\alpha t}}{1-\beta^2t^2}$$

12. $\mathcal{L}ogist\,(\alpha,\beta)$, **Logistic**:

$$f(x;\alpha,\beta) = \frac{e^{-(x-\alpha)/\beta}}{s\left(1+e^{-(x-\alpha)/\beta}\right)^2};\ \ F(x;\alpha,\beta) = \frac{1}{1+e^{-(x-\alpha)/\beta}}$$

$$\begin{aligned} \mu &= \alpha,\ \sigma^2 = \frac{1}{3}\pi^2\beta^2,\ \gamma_1 = 0,\ \gamma_2 = 1.2, \\ M_X(t) &= e^{\alpha t}B(1-\beta t, 1+\beta t)\ \text{ where } B \text{ is the beta function} \end{aligned}$$

13. $v\mathcal{M}(\alpha,\kappa)$, **von Mises**: $f(x;\alpha,\kappa) = \frac{1}{2\pi I_0(\kappa)}\exp(\kappa\cos(x-\alpha))\, I_{(-\pi,\pi)}(x)$, where $I_0(\kappa)$ is the modified Bessel function of order 0 and $\kappa>0$; $\mu=\tilde{x}=\alpha$;

$$\sigma^2 = 1-\frac{I_1(\kappa)}{I_0(\kappa)};\ cf = \varphi_X(\omega) = \frac{I_{|\omega|}(\kappa)}{I_0(\kappa)}e^{i\omega\alpha};$$

Some limits are

$$\lim_{\kappa\to 0} f(x;\alpha,\kappa) = \frac{1}{2\pi}I_{(-\pi,\pi)}(x);\ \lim_{\kappa\to\infty} f(x;\alpha,\kappa) = \frac{1}{\sqrt{2\pi/\kappa}}\exp\left\{-\frac{\kappa}{2}(x-\alpha)^2\right\}$$

$$\lim_{\kappa\to 0} v\mathcal{M}(\alpha,\kappa) = \mathcal{U}nif(-\pi,\pi);\ \lim_{\kappa\to\infty} v\mathcal{M}(\alpha,\kappa) = \mathcal{N}(\alpha,1/\kappa^2)$$

14. $\mathcal{P}ar(m,\alpha)$, **Pareto**: $f(x;m,\alpha) = \frac{\alpha m^\alpha}{x^{\alpha+1}}I_{(m,\infty)}(x)$, with m and α both positive.

$$\mu = \frac{m\alpha}{\alpha-1},\ \text{for } \alpha>1;\ \tilde{x} = m2^{1/\alpha};\ \sigma^2 = \frac{m^2\alpha}{(\alpha-1)(\alpha-2)},\ \text{for } \alpha>2$$

$$\mu_r' = \frac{m^n\alpha}{\alpha-n}\ \text{ for } n<\alpha$$

15. $\mathcal{E}xtr\,(\alpha,\beta)$, **Extreme Value**: $cdf = F\,(x;\alpha,\beta) = \exp\left\{-e^{-(x-\alpha)/\beta}\right\}$, for $\beta > 0$.

$$\mu = \alpha + \beta\gamma,\ \sigma^2 = \frac{1}{6}\pi\beta^2,\ \tilde{x} = \alpha - \beta\log(\log 2)\,,\ M_X\,(t) = e^{\alpha t}\Gamma\,(1-\beta t)\,,\text{ for } t < 1/\beta$$

16. t_n, t-**distribution**: $t_n = \dfrac{\mathcal{N}(0,1)}{\sqrt{\dfrac{\chi_n^2}{n}}}$ when numerator and denominator are independent and $n > 0$;

$$f(x;n) = \frac{\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)};\quad \mu = 0;\quad \sigma^2 = \frac{n}{n-2};\ \gamma_1 = 0 \text{ for } n = 4:\infty\text{, and } \gamma_2 = \frac{6}{n-4} \text{ for } n = 5:\infty$$

t_n only has moments up to order $n-1$, hence, the mgf *does not exist*.

17. $F_{m,n}$, F-**distribution**: $F_{m,n} = \dfrac{\chi_m^2/m}{\chi_n^2/n}$, when numerator and denominator are independent.

$$f(x;m,n) = \frac{m^{m/2}n^{n/2}}{B\left(\frac{m}{2},\frac{n}{2}\right)}x^{(m-2)/2}\,(n+mx)^{-(m+n)/2}\,I_{(0,\infty)}(x);$$

$$\mu = \frac{n}{n-2};\quad \sigma^2 = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

18. $\mathcal{N}_2(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$, **Bivariate normal**:

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2}Q\right\},$$

where

$$Q := \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

Then $X_1 \sim \mathcal{N}(\mu_1,\sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2,\sigma_2^2)$, $X|y \sim \mathcal{N}\left(\beta_x, \sigma_1^2(1-\rho^2)\right)$, and $Y|x \sim \mathcal{N}\left(\beta_y, \sigma_2^2(1-\rho^2)\right)$, where

$$\beta_x = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}\,(y-\mu_2) \text{ and } \beta_y = \mu_2 + \rho\frac{\sigma_2}{\sigma_2}\,(x-\mu_1)$$

Also

$$M_{X_1,X_2}(t_1,t_2) = \exp\left\{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}\left(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2\right)\right\}$$

19. $\chi_n'^2\,(\delta)$ **Noncentral chi-square** : If $Z_1, Z_2, \ldots Z_n$ are independent $\mathcal{N}\left(\mu_k,\sigma_k^2\right)$, then

$$\chi_n'^2\,(\delta) = \sum_{k=1}^{n}\left(\frac{X_k}{\sigma_k}\right)^2 \text{ where } \delta := \sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k}\right)^2 \text{ is the } \textbf{noncentrality parameter}$$

The pdf is

$$f\,(x;n,\delta) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{\lambda}{2}\right)^k e^{-\delta/2}\frac{x^{(n/2)+k-1}e^{-x/2}}{2^{(n/2)+k}\Gamma\left(\frac{n}{2}+k\right)}I_{(0,\infty)} \text{ for } \delta > 0,\ n = 1:\infty$$

$\mu = E\,(X) = n + \delta$, $\text{var}\,(X) = 2\,(n+2\delta)$, $M\,(t) = (1-2t)^{-n/2}\exp\left\{\delta t/\,(1-2t)\right\}$, $\kappa_r = 2^{r-1}\,(r-1)!\,(n+r\delta)$

$$\mu_r' = 2^r\Gamma\left(r+\frac{n}{2}\right)\sum_{k=0}^{\infty}\binom{r}{k}\frac{(\delta/2)^k}{\Gamma\left(k+\frac{n}{2}\right)}$$

20. $t_n'\,(\delta)$ **Noncentral** t: $t_n'\,(\delta) := \left(\mathcal{N}\,(\delta,1)\right)/\sqrt{\chi_n^2/n}$ and the pdf is

$$f\,(x;n) = \frac{n^{n/2}}{\Gamma\,(n/2)}\frac{e^{-\delta/2}}{\sqrt{\pi}\,(n+x^2)^{(n+1)/2}}\sum_{k=0}^{\infty}\Gamma\left(\frac{n+k+1}{2}\right)\left(\frac{\delta^k}{k!}\right)\left(\frac{2x^2}{n+x^2}\right)^{k/2}I_{(0,\infty)}$$

$$\mu = \delta\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{n}{2}} \text{ for } n > 1,\ \text{var}\,(X) = \frac{n\left(1+\delta^2\right)}{n-2} - \frac{\mu^2 n}{2}\frac{\Gamma^2\left((n-1)/2\right)}{\Gamma^2\,(n/2)} \text{ for } n > 2$$

21. $F'_{m,n}(\delta)$ **Noncentral** F: $F'_{m,n}(\delta) = \chi'^2_m(\delta)/\chi^2_n$ and the pdf is

$$f(x;m,n,\delta) = e^{-\delta/2}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{\delta}{2}\right)^k \frac{\Gamma\left(\frac{1}{2}(m+n+2k)\right)}{\Gamma\left(\frac{1}{2}(m+2k)\right)\Gamma\left(\frac{n}{2}\right)}\frac{x^{(m+2k)/-1}}{(1+x)^{(m+n+2k)/2}}I_{(0,\infty)}$$

$$\mu = \frac{(m+\delta)n}{(n-2)m} \text{ for } n>2, \ \operatorname{var}(X)^2 = \frac{(m+\delta)^2+2(m+\delta)n}{(n-2)(n-4)m^2} - \frac{(m+\delta)^2 n^{2^2}}{(n-2)^2 m} \text{ for } n>4$$

AdditionTheorems, Division Statements, Miscellaneous Relations

Each of the following sums are of independent rvs of the type indicated.

1. $\sum \mathcal{B}in(n_k,p) = \mathcal{B}in\left(\sum n_k, p\right)$
2. $\sum_1^n \mathcal{G}eo(p) = \mathcal{B}in^*(n,p)$
3. $\sum \mathcal{P}ois(\lambda_k) = \mathcal{P}ois\left(\sum \lambda_k\right)$
4. $\sum_1^n \mathcal{E}xp(\beta) = \mathcal{G}am(n,\beta)$
5. $\sum_1^n \mathcal{G}am(\alpha_k,\beta) = \mathcal{G}am(\sum \alpha_k, \beta)$
6. $\sum a_k \mathcal{N}(\mu_k,\sigma_k^2) = \mathcal{N}\left(\sum a_k\mu_k, \sum a_k^2\sigma_k^2\right)$
7. $\chi_1^2 = \{\mathcal{N}(0,1)\}^2$
8. $\sum \chi^2_{n_k} = \chi^2_{\Sigma n_k}$
9. $\sum \chi'^2_{n_k} = \chi'^2_{\Sigma n_k}$
10. $X_1,\ldots,X_n$ iid $\mathcal{N}(\mu,\sigma^2) \Leftrightarrow \bar{X} \sim \mathcal{N}(\mu,\frac{\sigma^2}{n})$ independent of $(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1)$
11. If X,Y iid $\mathcal{N}(0,1)$ then
    (a) $\frac{X}{|X|} \sim \mathcal{C}auchy(0,1) = t_1$
    (b) $\frac{X+Y}{X-Y} \sim \mathcal{C}auchy(0,1)$
    (c) $U = \frac{XY}{\sqrt{X^2+Y^2}} \sim \mathcal{N}\left(0,\frac{1}{\sqrt{2}}\right)$ independent of $V = \frac{X^2-Y^2}{X^2+Y^2}$ which is also normal.
12. $X_1,\ldots,X_n$ iid $\mathcal{U}nif(0,1) \Rightarrow X_{(r;n)} \sim \mathcal{B}eta(r,n-r+1)$
13. $X \sim \mathcal{U}nif(0,1) \Rightarrow -2\log X \sim \chi^2(2)$
14. $X_1,X_2,\ldots,X_n$ iid $\mathcal{E}xp(\beta) \Rightarrow X_{(1;n)} \sim \mathcal{E}xp(\beta/n)$
15. $X \sim \mathcal{G}am(\alpha,\beta) \Rightarrow \frac{2X}{\beta} \sim \chi^2_{2\alpha}$
16. $X \sim \chi^2_m$ independent of $Y \sim \chi^2_n \Rightarrow \frac{X}{X+Y} \sim \mathcal{B}eta\left(\frac{m}{2},\frac{n}{2}\right)$
17. $F \sim F_{m,n} \Rightarrow \frac{(m/n)F}{1+(m/n)F} \sim F_{m/2,n/2}$
18. $X \sim \mathcal{B}eta(\alpha_1,\beta_1)$ independent of $Y \sim \mathcal{B}eta(\alpha_2,\beta_2)$ $\begin{cases} \alpha_1 = \alpha_2+\beta_2 & \Rightarrow XY \sim \mathcal{B}eta(\alpha_2,\beta_1+\beta_2) \\ \alpha_2 = \alpha_1+\beta_1 & \Rightarrow XY \sim \mathcal{B}eta(\alpha_1,\beta_1+\beta_2) \end{cases}$
19. $(X,Y) \sim \mathcal{N}_2(0,0;1,1;\rho) \Rightarrow \frac{Y}{X} \sim \mathcal{C}auchy(0,1)$
20. $X \sim N\mathcal{B}in(r,p)$ and $Y \sim \mathcal{B}in(n,p) \Rightarrow \Pr(X \le n) = \Pr(Y \ge r)$. In terms of cdfs, this is $F_X(n;r,p) = 1 - F_Y(r;n,p)$
21. $X \sim \mathcal{G}am(n,\beta)$ and $Y \sim \mathcal{P}ois(1/\beta) \Rightarrow \Pr(X \le x;n,\beta) = 1 - \Pr(Y \le n-1;x,\beta)$. In terms of cdfs, this is $F_X(x;n,\beta) = 1 - F_Y(n-1;x,\beta)$

22. $X \sim \log\mathcal{N}(\alpha_1, \beta_1)$ and $Y \sim \log\mathcal{N}(\alpha_2, \beta_2)$ independent, then $XY \sim \log\mathcal{N}(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$ and $X/Y \sim \log\mathcal{N}(\alpha_1 - \alpha_2, \beta_1 - \beta_2)$

23. For any continuous rv X with cdf $F(x)$, the r^{th} order statistic $X_{(r;n)}$ has cdf $G_r(y) = H\left(F(y); r, n-r+1\right)$, where H is the cdf of a $\mathcal{B}eta(r, n-r+1)$ rv.

24. **Gamma function**: $\Gamma\left(\alpha+1\right) = \alpha\Gamma\left(\alpha\right) := \int_0^\infty t^\alpha e^{-t} dt$; $\Gamma\left(1\right) = 1$; $\Gamma\left(n+1\right) = n!$ when n is a nonnegative integer, and $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

25. $\int_0^\infty t^\alpha e^{-\beta t} dt = \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}}$, $\beta > 0$

26. **Incomplete gamma function**: $\gamma\left(a, x\right) := \int_0^x t^{a-1} e^{-t} dt$ for $a > 0$ and $x > 0$. $P\left(a, x\right) := \gamma\left(a, x\right) / \Gamma\left(a\right)$ is the cdf of the gamma distribution. The corresponding tail probability is

$$\frac{\Gamma\left(a, x\right)}{\Gamma\left(a\right)} := 1 - \frac{\gamma\left(a, x\right)}{\Gamma\left(a\right)} = \frac{1}{\Gamma\left(a\right)} \int_x^\infty t^{a-1} e^{-t} dt$$

27. **Beta function**: $B\left(\alpha, \beta\right) := \int_0^1 t^{\alpha-1} \left(1-t\right)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, for $\alpha > 0$, $\beta > 0$

28. **Incomplete beta function**: $B_x\left(\alpha, \beta\right) := \int_0^x t^{\alpha-1} \left(1-t\right)^{\beta-1} dt$. $I_x\left(\alpha, \beta\right) := B_x\left(\alpha, \beta\right) / B\left(\alpha, \beta\right)$ is the cdf of the beta distribution.