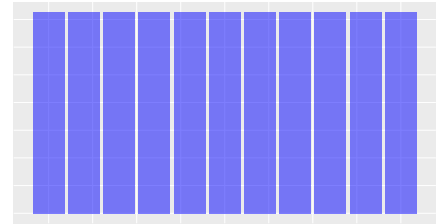


# Probability Distributions

## (Discrete) Uniform

Each integer value in a specified range is assigned the same value.

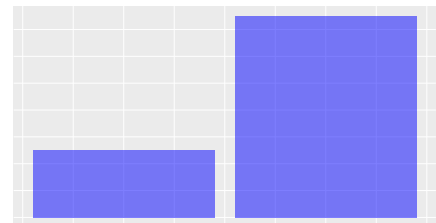
Parameters	lower bound $a \in \mathbb{Z}$ , upper bound $b \in \mathbb{Z}$ s.t. $a < b$
Possible values	$\{a, a + 1, \dots, b - 1, b\}$
Notation	$U(a, b)$
Probability function	$p(x) = \frac{1}{b-a+1}$
Expectation	$\frac{a+b}{2}$
Variance	$\frac{(b-a)(b-a+2)}{12}$
MGF	$M(t) = \frac{1}{b-a+1} \sum_{k=a}^b e^{kt}$
Kernel	1



## Bernoulli

Represents the success of a single experiment as a binary outcome.

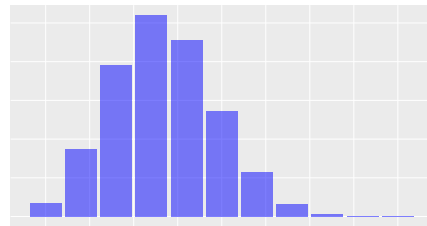
Parameters	probability of success $0 \leq p \leq 1$
Possible values	$\{0, 1\}$
Notation	Bern( $p$ )
Probability function	$p(x) = p^x(1-p)^{1-x}$
Expectation	$p$
Variance	$p(1-p)$
MGF	$M(t) = 1 + p(e^t - 1)$
Kernel	$p^x(1-p)^{-x}$



## Binomial

Represents the number of successes in a fixed number of independent and repeated trials of the same Bernoulli experiment.

Parameters	number of trials $n \in \mathbb{N}$ , probability of success on a single trial $0 \leq p \leq 1$
Possible values	$\{0, 1, 2, \dots, n\}$
Notation	$\text{Bin}(n, p)$
Probability function	$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$
Expectation	$np$
Variance	$np(1-p)$
MGF	$M(t) = [1 + p(e^t - 1)]^n$
Kernel	$\binom{n}{x} p^x (1-p)^{n-x}$



## Related Distributions

- If  $X \sim \text{Bin}(n, p)$ , then

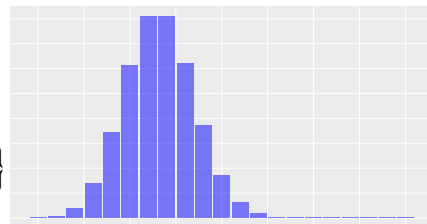
$$X = \sum_{i=1}^n X_i,$$

where the  $X_i \sim \text{Bern}(p)$  independently.

## Hypergeometric

Represents the number of when drawing a fixed number of samples from a population containing a known number of successes.

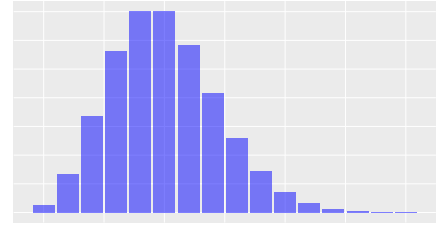
Possible values	$\{\max\{0, n - (n - (N - k))\}, \dots, \min\{n, k\}\}$ size of population $N \in \mathbb{N}$ ,
Parameters	number of successes in population $k \in \mathbb{N}$ , number of samples drawn $n \in \mathbb{N}$
Notation	Hypergeometric $(N, k, n)$
Probability function	$p(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$
Expectation	$\frac{nk}{N}$
Variance	$n \left( \frac{k}{N} \right) \left( \frac{N-k}{N} \right) \left( \frac{N-n}{N-1} \right)$
MGF	No useful expression
Kernel	$\binom{k}{x} \binom{N-k}{n-x}$



## Poisson

Represents the number of events occurring in a fixed interval.

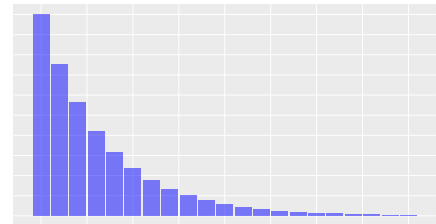
Possible values	$\{0, 1, 2, \dots\}$
Parameters	average number of events in interval $\lambda > 0$ ,
Notation	$\text{Po}(\lambda)$
Probability function	$p(x) = \frac{e^{-\lambda} \lambda^x}{x!}$
Expectation	$\lambda$
Variance	$\lambda$
MGF	$M(t) = e^{\lambda(e^t - 1)}$
Kernel	$\frac{\lambda^x}{x!}$



## Geometric

Represents the number of failed Bernoulli trials preceding the first success.

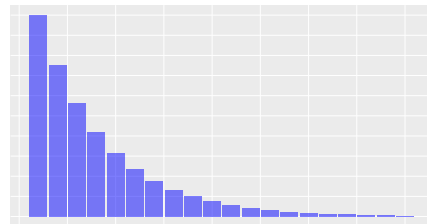
Parameters	probability of success on a single trial $0 \leq p \leq 1$ ,
Possible values	$\{0, 1, 2, \dots\}$
Notation	$\text{Geom}(p)$
Probability function	$p(x) = (1 - p)^x p$
Expectation	$\frac{1-p}{p}$
Variance	$\frac{1-p}{p^2}$
MGF	$M(t) = \frac{p}{1 - (1-p)e^t}$ for $t < -\ln(1-p)$
Kernel	$(1-p)^x$



## Geometric (alternative)

Represents the position of the first success in a sequence of Bernoulli trials.

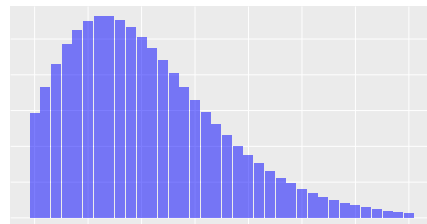
Parameters	probability of success on a single trial $0 \leq p \leq 1$ ,
Possible values	$\{1, 2, \dots\}$
Notation	$\text{Geom}(p)$
Probability function	$p(x) = (1 - p)^{x-1}p$
Expectation	$\frac{1}{p}$
Variance	$\frac{1-p}{p^2}$
MGF	$M(t) = \frac{pe^t}{1-(1-p)e^t}$ for $t < -\ln(1-p)$
Kernel	$(1-p)^x$



## Negative Binomial

Represents the number of Bernoulli trials preceding the  $r$ th success.

Parameters	number of desired successes $r \in \mathbb{N}$ , probability of success on a single trial $0 \leq p \leq 1$ ,
Possible values	$\{r, r+1, \dots\}$
Notation	$\text{NB}(r, p)$
Probability function	$p(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r$
Expectation	$\frac{r}{p}$
Variance	$\frac{r(1-p)}{p^2}$
MGF	$M(t) = \left[ \frac{pe^t}{1-(1-p)e^t} \right]^r$ for $t < -\ln p$
Kernel	$\binom{x-1}{r-1} (1-p)^x$



## Related Distributions

- If  $X \sim \text{NB}(r, p)$ , then

$$X = \sum_{i=1}^n X_i,$$

where the  $X_i \sim \text{Geom}(p)$  independently.

## Negative Binomial (Alternative)

Represents the number of failures preceding the  $r$ th success in a sequence of Bernoulli trials.

### Parameters

number of desired successes  $r \in \mathbb{N}$ ,  
probability of success on a single trial  
 $0 \leq p \leq 1$ ,

### Possible values

$\{0, 1, 2, \dots\}$

### Notation

$\text{NB}(r, p)$

### Probability function

$$p(x) = \binom{x+r-1}{x} (1-p)^x p^r$$

### Expectation

$$\frac{r(1-p)}{p}$$

### Variance

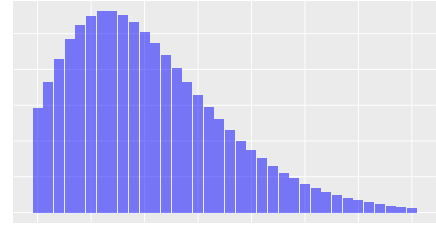
$$\frac{r(1-p)}{p^2}$$

### MGF

$$M(t) = \left[ \frac{p}{1-(1-p)e^t} \right]^r \text{ for } t < -\ln p$$

### Kernel

$$\binom{x+r-1}{x} (1-p)^x$$

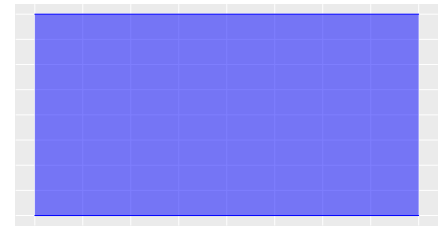


## Continuous

### (Continuous) Uniform

Each value in a specified interval has the same probability density.

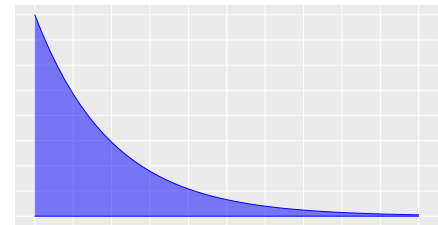
Parameters	lower bound $a \in \mathbb{R}$ , upper bound $b \in \mathbb{R}$ s.t. $a < b$
Possible values	$(a, b)$
Notation	$U(a, b)$
Probability density function	$f(x) = \frac{1}{b-a}$
Cumulative distribution function	$F(x) = \begin{cases} 0, & \text{if } x \leq a, \\ \frac{x-a}{b-a}, & \text{if } a < x < b, \\ 1, & \text{if } b \leq x. \end{cases}$
Expectation	$\frac{a+b}{2}$
Variance	$\frac{(b-a)^2}{12}$
MGF	$M(t) = \begin{cases} 1, & \text{if } t = 0, \\ \frac{e^{tb} - e^{ta}}{t(b-a)}, & \text{if } t \neq 0. \end{cases}$
Kernel	1



## Exponential

Represents the waiting time until an event occurs.

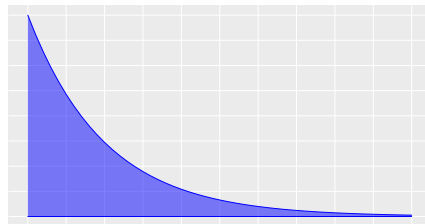
Parameters	rate of events $\lambda > 0$
Possible values	$(0, \infty)$
Notation	$\text{Exp}(\lambda)$
Probability density function	$f(x) = \lambda e^{-\lambda x}$
Cumulative distribution function	$F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{if } x > 0. \end{cases}$
Expectation	$\frac{1}{\lambda}$
Variance	$\frac{1}{\lambda^2}$
MGF	$M(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$
Kernel	$e^{-\lambda x}$



## Exponential (Alternative)

Characterised by average time until the first event  $\beta = 1/\lambda$ .

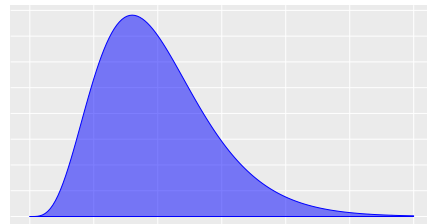
Parameters	average time until event $\beta > 0$
Possible values	$(0, \infty)$
Notation	$\text{Exp}(\beta)$
Probability density function	$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$
Cumulative distribution function	$F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1 - e^{-\frac{x}{\beta}}, & \text{if } x > 0. \end{cases}$
Expectation	$\beta$
Variance	$\beta^2$
MGF	$M(t) = \frac{1}{1-\beta t}$ for $t < \frac{1}{\beta}$
Kernel	$e^{-\frac{x}{\beta}}$



## Gamma

Generalisation of the exponential distribution. For integer shape parameter, represents the waiting time until the  $r$ th event.

Parameters	shape $r > 0$ scale $\lambda > 0$
Possible values	$(0, \infty)$
Notation	Gamma $(a, b)$
Probability density function	$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$
Cumulative distribution function	no simple expression
Expectation	$\frac{r}{\lambda}$
Variance	$\frac{r}{\lambda^2}$
MGF	$M(t) = \left(\frac{\lambda}{\lambda-t}\right)^r$ for $t < \lambda$
Kernel	$x^{r-1} e^{-\lambda x}$



## Properties

- If  $X \sim \text{Gamma}(r, \lambda)$ , then for any  $k > 0$ ,  $kX \sim \text{Gamma}\left(r, \frac{\lambda}{k}\right)$

## Related Distributions

- If  $r \in \mathbb{N}$  and  $X \sim \text{Gamma}(r, \lambda)$ , then

$$X = \sum_{i=1}^n X_i,$$

where  $X_i \sim \text{Exp}(\lambda)$  independently.

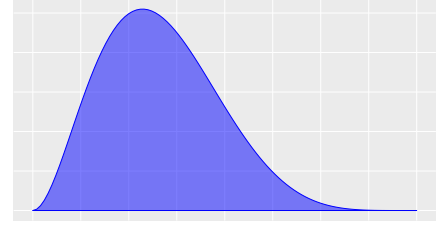
- If  $X \sim \chi_r^2$ , then  $X \sim \text{Gamma}\left(\frac{r}{2}, \frac{1}{2}\right)$ .



## Beta

Represents a random variable which is positive and bounded. Useful for modelling proportions and percentages.

Parameters	shape $\alpha > 0$ shape $\beta > 0$
Possible values	$(0, 1)$
Notation	Beta $(a, b)$
Probability density function	$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$
Cumulative distribution function	no simple expression
Expectation	$\frac{\alpha}{\alpha+\beta}$
Variance	$\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$
MGF	exists but no simple expression
Kernel	$x^{\alpha-1} (1-x)^{\beta-1}$



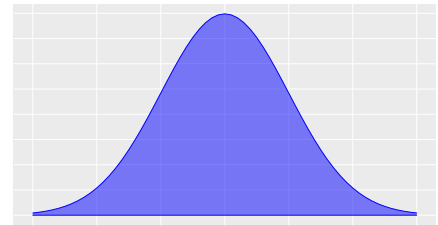
## Related Distributions

- If  $X \sim \text{Gamma}(\alpha, \lambda)$  and  $Y \sim \text{Gamma}(\beta, \lambda)$  independently, then  $\frac{X}{X+Y} \sim \text{Beta}(\alpha, \beta)$ .
- The Beta distribution can be generalised to any bounded interval  $[a, b]$ , with properties

## Normal

The classic bell-shaped curve.

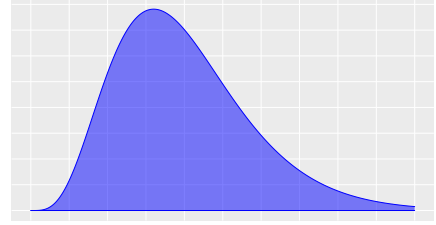
Parameters	mean $\mu \in \mathbb{R}$ variance $\sigma^2 > 0$
Possible values	$\mathbb{R}$
Notation	$\mathcal{N}(\mu, \sigma^2)$
Probability density function	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Cumulative distribution function	no simple expression
Expectation	$\mu$
Variance	$\sigma^2$
MGF	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$
Kernel	$e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



$\chi^2$

The sum of the squares of  $n$  independent standard normal random variables.

Parameters	degrees of freedom $n \in \mathbb{N}$
Possible values	$(0, \infty)$ if $n = 1$ ; $[0, \infty)$ otherwise
Notation	$\chi_n^2$
Probability density function	$f(x) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$
Cumulative distribution function	$F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \frac{1}{\Gamma(\frac{n}{2})} \int_0^{x/2} t^{\frac{n}{2}-1} e^{-t} dt. & \text{if } x > 0 \end{cases}$
Expectation	$n$
Variance	$2n$
MGF	$(1 - 2t)^{-\frac{n}{2}}$
Kernel	$x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$



### Related Distributions

- If  $X \sim \chi_n^2$ , then

$$X = \sum_{i=1}^n Z_i^2,$$

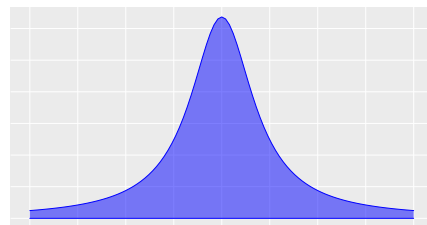
where  $Z_i \sim \mathcal{N}(0, 1)$  independently.

- If  $X \sim \chi_r^2$ , then  $X \sim \text{Gamma}(\frac{r}{2}, \frac{1}{2})$ .
- If  $Y \sim F_{\nu_1, \nu_2}$ , then  $\lim_{\nu_2 \rightarrow \infty} \nu_1 Y \sim \chi_{\nu_1}^2$ .

## Cauchy

The ratio of two independent, identically distributed and central normal random variables.

The sum of the squares of  $n$  independent standard normal random variables.



Parameters

location  $\mu \in \mathbb{R}$

scale  $\sigma > 0$

Possible values

$\mathbb{R}$

Notation

Cauchy  $(\mu, \sigma)$

Probability density function

$$f(x) = \frac{1}{\pi \sigma \left(1 + \frac{2(x-\mu)^2}{\sigma^2}\right)}$$

Cumulative distribution function

$$F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\sigma}\right), & \text{if } x > 0 \end{cases}$$

Expectation

does not exist

Variance

does not exist

MGF

does not exist

Kernel

$$\sigma + 2(x - \mu)$$

## Related Distributions

- If  $Z_1, Z_2 \sim \mathcal{N}(0, 1)$  independently, then  $Z_1/Z_2 \sim \text{Cauchy}(0, 1)$ .
- If  $X \sim \text{Cauchy}(\mu, \sigma)$  then  $X \sim t_n(\mu, \sigma)$ .

$t$

The ratio of a standard normal random variable and  $\chi^2$  random variable.

Parameters

location  $\mu \in \mathbb{R}$

scale  $\sigma > 0$

degrees of freedom  $\nu \in \mathbb{N}$

Possible values

$(0, \infty)$  if  $n = 1$ ;  $[0, \infty)$  otherwise

Notation

$t_n(\mu, \sigma^2)$

Probability density function

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sigma\sqrt{\pi\nu}} \left[ 1 + \frac{1}{\nu} \left( \frac{x-\mu}{\sigma} \right)^2 \right]^{-\frac{\nu+1}{2}}$$

Cumulative distribution function

no simple expression

Expectation

$\mu$  for  $\nu \geq 2$

Variance

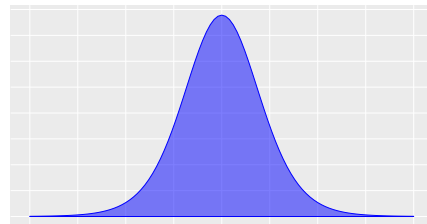
$\frac{\sigma^2\nu}{\nu-2}$  for  $\nu \leq 3$

MGF

does not exist

Kernel

$$\left[ 1 + \frac{1}{\nu} \left( \frac{x-\mu}{\sigma} \right)^2 \right]^{-\frac{\nu+1}{2}}$$



**Related Distributions**

- If  $Z \sim \mathcal{N}(0, 1)$  and  $W \sim \chi^2_\nu$  independently, then

$$\mu + \sigma \frac{Z}{\sqrt{\frac{W}{\nu}}} \sim t_\nu(\mu, \sigma^2).$$

$F$

The ratio of the sample variance of two normal random variables.

Parameters

degrees of freedom  $\nu_1 \in \mathbb{N}$

degrees of freedom  $\nu_2 \in \mathbb{N}$

Possible values

$(0, \infty)$

Notation

$F_n \mu, \sigma^2$

Probability density function

$$f(x) = \frac{\Gamma(\frac{\nu_1 \nu_2}{2}) \nu_1^{\nu_1/2} \nu_2^{\nu_2/2} x^{\nu_1/2-1}}{\Gamma(\frac{\nu_1}{2}) \Gamma(\frac{\nu_2}{2}) \Gamma(\frac{\nu_1 + \nu_2}{2}) 2^{\frac{\nu_1 + \nu_2}{2}}}$$

Cumulative distribution function

no simple expression

Expectation

$\frac{\nu_2}{\nu_2 - 2}$  for  $\nu \geq 2$

Variance

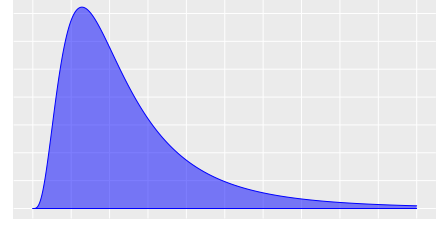
$\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$  for  $\nu \geq 4$

MGF

does not exist

Kernel

$$\frac{x^{\nu_1/2-1}}{(\nu_1 x + \nu_2)^{\frac{\nu_1 + \nu_2}{2}}}$$



## Related Distributions

- If  $U \sim \chi_{\nu_1}^2$  and  $V \sim \chi_{\nu_2}^2$  independently, then

$$\frac{U/\nu_1}{V/\nu_2} \sim F_{\nu_1, \nu_2}.$$

- If  $X_1 \sim \text{Gamma}(\alpha_1, \beta_1)$  and  $X_2 \sim \text{Gamma}(\alpha_2, \beta_2)$  independently, then

$$\frac{\alpha_1 \beta_1 X_1}{\alpha_2 \beta_2 X_2} \sim F_{2\alpha_1, 2\alpha_2}.$$

- If  $X \sim \text{Beta}(\alpha/2, \beta/2)$  then

$$\frac{\beta X}{\alpha(1-X)} \sim F_{\alpha, \beta}.$$

- If  $X \sim t_\nu$ , then  $X^2 \sim F_{1, \nu}$ .

# Multivariate

## Discrete

### Trinomial

A generalisation of the binomial distribution to 3 possible outcomes.

Parameters	number of trials $n \in \mathbb{N}$ probabilities of success $\pi_1, \pi_2 > 0$ such that $\pi_1 + \pi_2 < 1$
Possible values	$x_1, x_2 \in \{0, 1, \dots, n\}$ such that $x_1 + x_2 \leq n$
Notation	$\text{Tri}(n, \pi_1, \pi_2)$
Probability function	$p(x_1, x_2) = \frac{n!}{x_1!x_2!(n-x_1-x_2)!} \pi_1^{x_1} \pi_2^{x_2} (1 - \pi_1 - \pi_2)^{n-x_1-x_2}$
Expectation	$\mathbb{E}[X_i] = n\pi_i$
Variance	$\text{var}(X_i) = n\pi_i(1 - \pi_i)$
Covariance	$\text{cov}(X_1, X_2) = -n\pi_1\pi_2$
Marginal	$X_1 \sim \text{Bin}(n, \pi_1)$ $X_2 \sim \text{Bin}(n, \pi_2)$
Conditional	$X_1   X_2 = x_2 \sim \text{Bin}\left(n - x_2, \frac{\pi_1}{1 - \pi_2}\right)$ $X_2   X_1 = x_1 \sim \text{Bin}\left(n - x_1, \frac{\pi_2}{1 - \pi_1}\right)$

### Multinomial

Generalisation of the binomial distribution to  $r$  possible outcomes.

Parameters	number of trials $n \in \mathbb{N}$ probabilities of success $\boldsymbol{\pi} = (\pi_1, \dots, \pi_r)$ such that $\pi_i \geq 0$ and $\sum_{i=1}^r \pi_i = 1$
Possible values	$x_1, \dots, x_r \in \{0, 1, \dots, n\}$ such that $\sum_{i=1}^r x_i = n$
Notation	$\text{Multi}_r(n, \boldsymbol{\pi})$
Probability function	$p(x_1, \dots, x_r) = \frac{n!}{x_1! \dots x_r!} \pi_1^{x_1} \dots \pi_r^{x_r}$
Expectation	$\mathbb{E}[X_i] = n\pi_i$
Variance	$\text{var}(X_i) = n\pi_i(1 - \pi_i)$
Covariance	$\text{cov}(X_i, X_j) = -n\pi_i\pi_j$ for $i \neq j$

## Continuous

### Uniform on Unit Square

A bivariate extension of the univariate (continuous) uniform distribution.

Parameters	none
Possible values	$0 < x_1, x_2 < 1$
Probability density function	$f(x_1, x_2) = 1$

## Uniform on Unit Circle

Another bivariate extension of the univariate (continuous) uniform distribution.

Parameters	none
Possible values	$-1 < x_1, x_2 < 1$ such that $x_1^2 + x_2^2 < 1$
Probability density function	$f(x_1, x_2) = \frac{1}{\pi}$

## Dirichlet

Generalisation of the Beta distribution.

Parameters	$\alpha_1, \dots, \alpha_{r+1} > 0$
Possible values	$x_1, \dots, x_r > 0$ such that $\sum_{i=1}^r x_i < 1$
Notation	$\text{Dir}(\boldsymbol{\alpha})$
Probability density function	$f(x_1, \dots, x_r) = \frac{\Gamma(\alpha_1 + \dots + \alpha_{r+1})}{\Gamma(\alpha_1) \dots \Gamma(\alpha_{r+1})} \prod_{i=1}^r x_i^{\alpha_i - 1}$
Expectation	$\mathbb{E}[X_i] = \frac{\alpha_i}{\alpha_0}$
Variance	$\text{var}(X_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}$
Covariance	$\text{cov}(X_1, X_2) = -\frac{\alpha_1 \alpha_2}{\alpha_0^2(\alpha_0 + 1)}$ where $\alpha_0 = \sum_{i=1}^{r+1} \alpha_i$
Marginal	$X_i \sim \text{Beta}(\alpha_i, \alpha_0 - \alpha_i)$

## Multivariate Normal

Generalisation of the normal distribution.

Parameters	mean $\boldsymbol{\mu} \in \mathbb{R}^r$ non-negative definite variance matrix $\Sigma$
Possible values	$\mathbf{x} \in \mathbb{R}^r$
Notation	$\mathcal{N}_r(\boldsymbol{\mu}, \Sigma)$
Probability density function	$p(\mathbf{x}) = (2\pi)^{-\frac{r}{2}}  \Sigma ^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})}$
Expectation	$\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$
Variance	$\text{var}(\mathbf{X}) = \Sigma$
MGF	$e^{\mathbf{t}^T \boldsymbol{\mu} + \frac{1}{2} \mathbf{t}^T \Sigma \mathbf{t}}$