

Some common discrete distributions

Distribution	Notation	pmf $f(x)$	Range of X	Parameter range	$\mathbb{E}(X)$	$\text{Var}(X)$	$\mathbb{E}(z^X)$
Discrete uniform	$X \sim U\{1, \dots, n\}$	$\frac{1}{n}$	$\{1, \dots, n\}$	$n \in \mathbb{N}$	$\frac{1}{2}(n+1)$	$\frac{1}{12}(n^2-1)$	$\frac{1}{n} \sum_{i=1}^n z^i$
Binomial	$X \sim \text{Bin}(n, p)$	$\binom{n}{x} p^x (1-p)^{n-x}$	$\{0, 1, \dots, n\}$	$n \in \mathbb{N}, p \in [0, 1]$	np	$np(1-p)$	$\{pz + (1-p)\}^n$
Poisson	$X \sim \text{Poi}(\lambda)$	$e^{-\lambda} \frac{\lambda^x}{x!}$	$\{0, 1, \dots\}$	$\lambda \in [0, \infty)$	λ	λ	$e^{\lambda(z-1)}$
Negative binomial	$X \sim \text{NegBin}(k, p)$	$\binom{x-1}{k-1} p^k (1-p)^{x-k}$	$\{k, k+1, \dots\}$	$k \in \mathbb{N}, p \in [0, 1]$	$\frac{k}{p}$	$\frac{k(1-p)}{p^2}$	$\frac{(pz)^k}{\{1-(1-p)z\}^k}$
Multinomial	$X \sim \text{Multi}(n, p_1, \dots, p_k)$	$\frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k}$	$(n_1, \dots, n_k) \in \{0, 1, \dots, n\}^k : \sum_i n_i = n$	$p_1, \dots, p_k \in [0, 1] : \sum_i p_i = 1, n \in \mathbb{N}$	$(np_1, \dots, np_k)$	$\text{Cov}(X_i, X_j) = \begin{cases} np_i(1-p_i) & i=j \\ -np_i p_j & i \neq j \end{cases}$	$\mathbb{E}(z_1^{X_1} \dots z_k^{X_k}) = (\sum_{i=1}^k p_i z_i)^n$

Notes:

1. The $\text{Bin}(1, p)$ distribution is also called the Bernoulli(p) distribution. If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$, then $\sum_{i=1}^n X_i \sim \text{Bin}(n, p)$. The $\text{Bin}(n, p)$ distribution models the number of successes in n independent trials, each with probability p of success.
2. The $\text{NegBin}(1, p)$ distribution is also called the Geometric(p) distribution. If $X_1, \dots, X_k \stackrel{\text{iid}}{\sim} \text{Geometric}(p)$, then $\sum_{i=1}^k X_i \sim \text{NegBin}(k, p)$. The $\text{NegBin}(k, p)$ distribution models the number of independent trials required to attain k successes, each with probability p of success.
3. The $\text{Multi}(n, p_1, \dots, p_k)$ distribution models the number of balls that appear in each of k buckets, when n balls are placed independently in the buckets and a ball falls in the i th bucket with probability p_i .

Some common absolutely continuous distributions

Distribution	Notation	pdf $f(x)$	Range	Parameter range	$\mathbb{E}(X)$	$\text{Var}(X)$	$\mathbb{E}(e^{tX})$
Uniform	$X \sim U[a, b]$	$\frac{1}{b-a}$	$[a, b]$	$(a, b) \in \mathbb{R}^2, a < b$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$	$\frac{e^{bt}-e^{at}}{t(b-a)}$
Normal	$X \sim N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	$\mathbb{R}$	$\mu \in \mathbb{R}, \sigma \in (0, \infty)$	μ	σ^2	$e^{t\mu + \sigma^2 t^2/2}$
Gamma	$X \sim \Gamma(\alpha, \lambda)$	$\frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$	$(0, \infty)$	$\alpha \in (0, \infty), \lambda \in (0, \infty)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\begin{cases} (\frac{\lambda}{\lambda-t})^\alpha & \text{if } t < \lambda \\ \infty & \text{if } t \geq \lambda \end{cases}$
Beta	$X \sim \text{Beta}(a, b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$	$(0, 1)$	$a \in (0, \infty), b \in (0, \infty)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	
Cauchy	$X \sim \text{Cauchy}$	$\frac{1}{\pi(1+x^2)}$	$\mathbb{R}$		Does not exist	∞	$\begin{cases} 1 & \text{if } t = 0 \\ \infty & \text{if } t \neq 0 \end{cases}$
Multivariate normal	$X \sim N_d(\mu, \Sigma)$	$\frac{e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}}{(2\pi)^{d/2} (\det \Sigma)^{1/2}}$	$\mathbb{R}^d$	$\mu \in \mathbb{R}^d, \Sigma \text{ pos. def.}$	μ	$\text{Cov}(X_i, X_j) = \Sigma_{ij}$	$\mathbb{E}(e^{t^T X}) = e^{t^T \mu + t^T \Sigma t/2}$

Notes:

1. The $\Gamma(1, \lambda)$ distribution is the same as the $\text{Exp}(\lambda)$ distribution. If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exp}(\lambda)$, then $\sum_{i=1}^n X_i \sim \Gamma(n, \lambda)$.
2. For $n \in \mathbb{N}$, the $\Gamma(\frac{n}{2}, \frac{1}{2})$ distribution is the same as the χ_n^2 distribution. If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(0, 1)$, then $\sum_{i=1}^n X_i^2 \sim \chi_n^2$.
3. Recall that the Gamma function is defined, for $z \in \mathbb{C}$ with $\text{Re}(z) > 0$, by $\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} dt$. If $n \in \mathbb{N}$, then $\Gamma(n) = (n-1)!$. The function $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is called the beta function.
4. We can also define the degenerate normal distribution: say $X \sim N(\mu, 0)$ if $\mathbb{P}(X = \mu) = 1$. Then we say $X = (X_1, \dots, X_d)^\top \sim N_d(\mu, \Sigma)$ if every linear combination $t_1 X_1 + \dots + t_d X_d$ has a (possibly degenerate) univariate normal distribution. This more general definition includes situations like the following: let $X_1 \sim N(0, 1)$, and let $X := (X_1, X_1)^\top$. Then $X \sim N_2(0, \Sigma)$, where $\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. In this case, $\det \Sigma = 0$.