Appendix: Discrete Random Variables

1. Binomial distribution

Notation: $X \sim \text{Binomial}(n, p)$.

Description: number of successes in n independent trials, each with probability p of success.

Probability function:

$$f_X(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, \ldots, n.$$

Mean: $\mathbb{E}(X) = np$.

Variance: $\text{Var}(X) = np(1-p) = npq$, where $q = 1 - p$.

Sum: If $X \sim \text{Binomial}(n, p)$, $Y \sim \text{Binomial}(m, p)$, and X and Y are **independent**, then

$$X + Y \sim \text{Bin}(n + m,\ p).$$

---

2. Poisson distribution

Notation: $X \sim \text{Poisson}(\lambda)$.

Description: arises out of the Poisson process as the number of events in a fixed time or space, when events occur at a constant average rate. Also used in many other situations.

Probability function: $f_X(x) = \mathbb{P}(X = x) = \dfrac{\lambda^x}{x!} e^{-\lambda}$ for $x = 0, 1, 2, \ldots$

Mean: $\mathbb{E}(X) = \lambda$.

Variance: $\text{Var}(X) = \lambda$.

Sum: If $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Poisson}(\mu)$, and X and Y are **independent**, then

$$X + Y \sim \text{Poisson}(\lambda + \mu).$$

---

3. Geometric distribution

Notation: $X \sim \text{Geometric}(p)$.

Description: number of failures before the **first** success in a sequence of independent trials, each with $\mathbb{P}(\text{success}) = p$.

Probability function: $f_X(x) = \mathbb{P}(X = x) = (1-p)^x p \quad \text{for} \quad x = 0, 1, 2, \ldots$

Mean: $\mathbb{E}(X) = \dfrac{1-p}{p} = \dfrac{q}{p}$, where $q = 1 - p$.

Variance: $\text{Var}(X) = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$, where $q = 1 - p$.

Sum: if $X_1, \ldots, X_k$ are **independent**, and each $X_i \sim \text{Geometric}(p)$, then

$$X_1 + \ldots + X_k \sim \text{Negative Binomial}(k, p).$$

---

4. Negative Binomial distribution

Notation: $X \sim \text{NegBin}(k, p)$.

Description: number of failures before the ***k*th** success in a sequence of independent trials, each with $\mathbb{P}(\text{success}) = p$.

Probability function:

$$f_X(x) = \mathbb{P}(X = x) = \binom{k+x-1}{x} p^k (1-p)^x \quad \text{for} \quad x = 0, 1, 2, \ldots$$

Mean: $\mathbb{E}(X) = \dfrac{k(1-p)}{p} = \dfrac{kq}{p}$, where $q = 1 - p$.

Variance: $\text{Var}(X) = \dfrac{k(1-p)}{p^2} = \dfrac{kq}{p^2}$, where $q = 1 - p$.

Sum: If $X \sim \text{NegBin}(k,\, p)$, $Y \sim \text{NegBin}(m,\, p)$, and X and Y are **independent**, then

$$X + Y \sim \text{NegBin}(k + m,\, p).$$

---

5. Hypergeometric distribution

Notation: $X \sim \text{Hypergeometric}(N, M, n)$.

Description: Sampling without replacement from a finite population. Given N objects, of which M are 'special'. Draw n objects without replacement. X is the number of the n objects that are 'special'.

Probability function:

$$f_X(x) = \mathbb{P}(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad \text{for} \quad \begin{cases} x = \max(0,\ n + M - N) \\ \text{to}\ \ x = \min(n,\ M). \end{cases}$$

Mean: $\mathbb{E}(X) = np$, where $p = \dfrac{M}{N}$.

Variance: $\text{Var}(X) = np(1-p)\left(\dfrac{N-n}{N-1}\right)$, where $p = \dfrac{M}{N}$.

---

6. Multinomial distribution

Notation: $\boldsymbol{X} = (X_1, \ldots, X_k) \sim \text{Multinomial}(n;\ p_1, p_2, \ldots, p_k)$.

Description: there are n independent trials, each with k possible outcomes. Let $p_i = \mathbb{P}(\text{outcome } i)$ for $i = 1, \ldots k$. Then $\boldsymbol{X} = (X_1, \ldots, X_k)$, where X_i is the number of trials with outcome i, for $i = 1, \ldots, k$.

Probability function:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

$$\text{for } x_i \in \{0, \ldots, n\}\ \forall_i \text{ with } \sum_{i=1}^{k} x_i = n, \text{ and where } p_i \geq 0\ \forall_i,\ \sum_{i=1}^{k} p_i = 1.$$

Marginal distributions: $X_i \sim \text{Binomial}(n, p_i)$ for $i = 1, \ldots, k$.

Mean: $\mathbb{E}(X_i) = np_i$ for $i = 1, \ldots, k$.

Variance: $\text{Var}(X_i) = np_i(1-p_i)$, for $i = 1, \ldots, k$.

Covariance: $\text{cov}(X_i, X_j) = -np_ip_j$, for all $i \neq j$.

Continuous Random Variables

1. Uniform distribution

Notation: $X \sim \text{Uniform}(a, b)$.

Probability density function (pdf): $f_X(x) = \dfrac{1}{b-a}$ for $a < x < b$.

Cumulative distribution function:

$$F_X(x) = \mathbb{P}(X \leq x) = \frac{x-a}{b-a} \quad \text{for } a < x < b.$$

$$F_X(x) = 0 \text{ for } x \leq a, \text{ and } F_X(x) = 1 \text{ for } x \geq b.$$

Mean: $\mathbb{E}(X) = \dfrac{a+b}{2}$.

Variance: $\text{Var}(X) = \dfrac{(b-a)^2}{12}$.

---

2. Exponential distribution

Notation: $X \sim \text{Exponential}(\lambda)$.

Probability density function (pdf): $f_X(x) = \lambda e^{-\lambda x}$ for $0 < x < \infty$.

Cumulative distribution function:

$$F_X(x) = \mathbb{P}(X \leq x) = 1 - e^{-\lambda x} \quad \text{for } 0 < x < \infty.$$

$$F_X(x) = 0 \text{ for } x \leq 0.$$

Mean: $\mathbb{E}(X) = \dfrac{1}{\lambda}$.

Variance: $\text{Var}(X) = \dfrac{1}{\lambda^2}$.

Sum: if $X_1, \ldots, X_k$ are **independent**, and each $X_i \sim \text{Exponential}(\lambda)$, then

$$X_1 + \ldots + X_k \sim \text{Gamma}(k, \lambda).$$

---

3. Gamma distribution

Notation: $X \sim \text{Gamma}(k,\ \lambda)$.

Probability density function (pdf):

$$f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} \quad \text{for } 0 < x < \infty,$$

where $\Gamma(k) = \int_0^\infty y^{k-1} e^{-y}\, dy$ (the Gamma function).

Cumulative distribution function: no closed form.

Mean: $\mathbb{E}(X) = \dfrac{k}{\lambda}$.

Variance: $\text{Var}(X) = \dfrac{k}{\lambda^2}$.

Sum: if $X_1, \ldots, X_n$ are **independent**, and $X_i \sim \text{Gamma}(k_i,\ \lambda)$, then

$$X_1 + \ldots + X_n \sim \text{Gamma}(k_1 + \ldots + k_n,\ \lambda).$$

---

4. Normal distribution

Notation: $X \sim \text{Normal}(\mu,\ \sigma^2)$.

Probability density function (pdf):

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\{-(x-\mu)^2/2\sigma^2\}} \quad \text{for } -\infty < x < \infty.$$

Cumulative distribution function: no closed form.

Mean: $\mathbb{E}(X) = \mu$.

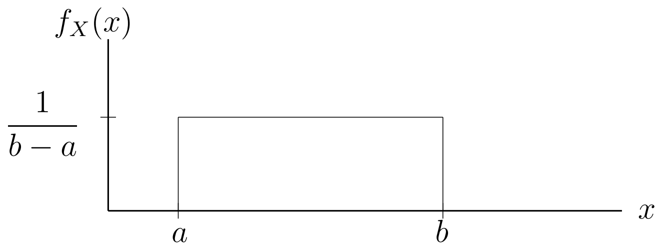
Variance: $\text{Var}(X) = \sigma^2$.

Sum: if $X_1, \ldots, X_n$ are **independent**, and $X_i \sim \text{Normal}(\mu_i,\ \sigma_i^2)$, then

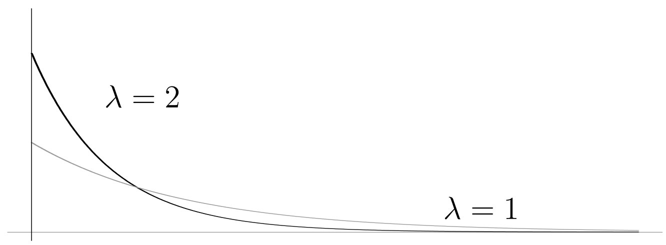
$$X_1 + \ldots + X_n \sim \text{Normal}(\mu_1 + \ldots + \mu_n,\ \sigma_1^2 + \ldots + \sigma_n^2).$$

Probability Density Functions

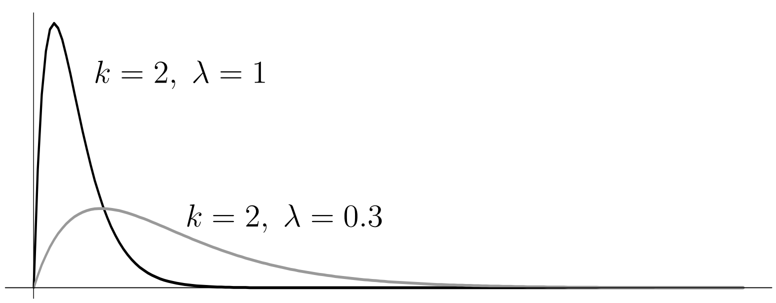
Uniform(a, b)



Exponential(λ)



Gamma(k, λ)



Normal(μ, σ^2)