

Statistics for Data Science - 2

Important results

Discrete random variables:

| Distribution | PMF ($f_X(k)$) | CDF ($F_X(x)$) | $E[X]$ | $\text{Var}(X)$ |
|--|---|--|-----------------|--------------------|
| Uniform(A) $A = \{a, a+1, \dots, b\}$ | $\frac{1}{n}, \quad x = k$ $n = b - a + 1$ $k = a, a+1, \dots, b$ | $\begin{cases} 0 & x < 0 \\ \frac{k-a+1}{n} & k \leq x < k+1 \\ & k = a, a+1, \dots, b-1, b \\ 1 & x \geq n \end{cases}$ | $\frac{a+b}{2}$ | $\frac{n^2-1}{12}$ |
| Bernoulli(p) | $\begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$ | $\begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$ | p | $p(1-p)$ |
| Binomial(n, p) | ${}^nC_k p^k (1-p)^{n-k},$ $k = 0, 1, \dots, n$ | $\begin{cases} 0 & x < 0 \\ \sum_{i=0}^k {}^nC_i p^i (1-p)^{n-i} & k \leq x < k+1 \\ & k = 0, 1, \dots, n \\ 1 & x \geq n \end{cases}$ | np | $np(1-p)$ |
| Geometric(p) | $(1-p)^{k-1} p,$ $k = 1, \dots, \infty$ | $\begin{cases} 0 & x < 0 \\ 1 - (1-p)^k & k \leq x < k+1 \\ & k = 1, \dots, \infty \end{cases}$ | $\frac{1}{p}$ | $\frac{1-p}{p^2}$ |
| Poisson(λ) | $\frac{e^{-\lambda} \lambda^k}{k!},$ $k = 0, 1, \dots, \infty$ | $\begin{cases} 0 & x < 0 \\ e^{-\lambda} \sum_{i=0}^k \frac{\lambda^i}{i!} & k \leq x < k+1 \\ & k = 0, 1, \dots, \infty \end{cases}$ | λ | λ |

Continuous random variables:

| Distribution | PDF ($f_X(k)$) | CDF ($F_X(x)$) | $E[X]$ | $\text{Var}(X)$ |
|---------------------------|--|---|---------------------|-----------------------|
| Uniform $[a, b]$ | $\frac{1}{b-a}, a \leq x \leq b$ | $\begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ |
| Exp(λ) | $\lambda e^{-\lambda x}, x > 0$ | $\begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$ | $\frac{1}{\lambda}$ | $\frac{1}{\lambda^2}$ |
| Normal(μ, σ^2) | $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right),$ $-\infty < x < \infty$ | No closed form | μ | σ^2 |

1. **Markov's inequality:** Let X be a discrete random variable taking non-negative values with a finite mean μ . Then,

$$P(X \geq c) \leq \frac{\mu}{c}$$

2. **Chebyshev's inequality:** Let X be a discrete random variable with a finite mean μ and a finite variance σ^2 . Then,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$