1 Table of Distributions

Notation	PMF/PDF and Support	Expected Value	Variance	MGF
Bernoulli Be(p)	$P(X=1)=p$ $P(X=0)=q$	p	pq	$q+pe^t$
Binomial Bin(n,p)	$P(X=k)=\binom{n}{k}p^kq^{n-k}$ $k=0,1,2,\ldots,n$	np	npq	$(q+pe^t)^n$
Geometric Geom(p)	$P(X=k)=pq^{k-1}$ $k=1,2,\ldots$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$, $qe^t<1$
Negative Binomial NB(r,p)	$P(X=k)=\binom{k-1}{r-1}p^rq^{k-r}$ $k=r,r+1,\ldots$	$\frac{r}{p}$	$\frac{rq}{p^2}$	$\left(\frac{pe^t}{1-qe^t}\right)^r$, $qe^t<1$
Hypergeometric HGeom(n,N,m)	$P(X=k)=\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$ $k=0,1,\ldots n$	$\frac{nm}{N}$	$\frac{nm}{N}\left[\frac{(n-1)(m-1)}{N-1}+1-\frac{nm}{N}\right]$	messy
Poisson Poisson(λ)	$P(X=k)=\frac{e^{-\lambda}\lambda^k}{k!}$ $k=0,1,2,\ldots$	λ	λ	$e^{\lambda(e^t-1)}$
Uniform U(a,b)	$f(x)=\frac{1}{b-a}$ $x\in(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Normal $\mathcal{N}(\mu,\sigma^2)$	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x\in(-\infty,\infty)$	μ	σ^2	$\exp\left(\mu t+\frac{\sigma^2t^2}{2}\right)$
Exponential Exp(λ)	$f(x)=\lambda e^{-\lambda x}$ $x\in[0,\infty)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$, $t<\lambda$
Gamma Gamma(a,λ)	$f(x)=\frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}$ $x\in[0,\infty)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)^a$, $t<\lambda$
Beta Beta(a,b)	$f(x)=\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}$ $x\in(0,1)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	messy
Chi-Square χ^2_n	$f(x)=\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}$ $x\in(0,\infty)$	n	$2n$	$(1-2t)^{-n/2}$ $t<1/2$

2 Useful Results

2.1 Probability

Sample space is the set of all possible outcomes of an experiment, usually denoted by S. For tossing two dice,

$$S=\{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(6,6)\}$$
$$=\{(i,j):1\le i,j\le 6\}$$

Event Any subset A of the sample space is an event

Odds The odds of an event A is defined by $\frac{P(A)}{1-P(A)}$

Increasing/Decreasing Events

If $\{E_n\}$ is an increasing sequence of events, then

$$\lim_{n\to\infty}E_n=\bigcup_{i=1}^{\infty}E_i$$

If $\{E_n\}$ is a decreasing sequence of events, then

$$\lim_{n\to\infty}E_n=\bigcap_{i=1}^{\infty}E_i$$

If $\{E_n\}$ is either an increasing or decreasing sequence of events, then

$$P(\lim_{n\to\infty}E_n)=\lim_{n\to\infty}P(E_n)$$

Conditional Probability

$$P(A|B)=P(A|BC)P(C|B)+P(A|BC^C)P(C^C|B)$$

Inclusion/Exclusion Principle Let $E_1, E_2, \ldots, E_n$ be any events, then

$$\begin{aligned}P(E_1 \cup E_2 \cup \cdots \cup E_n) = &\sum_{i=1}^{n} P(E_i) - \sum_{1 \le i_1 < i_2 \le n} P(E_{i_1} \cap E_{i_2}) + \cdots \\ &(-1)^{r+1} \sum_{1 \le i_1 < \cdots < i_r \le n} P(E_{i_1} \cap \cdots \cap E_{i_r}) \\ &+ \cdots + (-1)^{n+1} P(E_1 \cap \cdots \cap E_n)\end{aligned}$$

Derangement/Matching Problem

$$P(\text{at least one match}) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots (-1)^{n+1} \frac{1}{n!}$$

$$P(k \text{ matches}) = \binom{n}{k} \frac{1}{n(n-1)\cdots(n-k+1)} \cdot \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots (-1)^{n-k} \frac{1}{(n-k)!}\right)$$

2.2 Discrete RV

Suppose $X \sim$ Poisson(λ) and $Y \sim$ Bin(X, p), then $X - Y$ and Y are independent and

$$X - Y \sim \text{Poisson}(\lambda(1-p)), Y \sim \text{Poisson}(\lambda p)$$

2.3 Continuous RV

Exponential Distribution

$$P(X > s + t | X > s) = P(X > t)$$

Gamma Distribution

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha - 1} dt$$

- $\Gamma(1) = 1$ and $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- $\Gamma(n) = (n-1)!$ for $n = 1, 2, 3, \ldots$
- If $X_i \sim$ Exp(λ) independently, then $\sum_{i=1}^n X_i \sim$ Gamma(n, λ)

Beta Distribution

$$\begin{aligned}B(a, b) &= \int_0^1 t^{a-1}(1-t)^{b-1} dt \\ &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\end{aligned}$$

$$\text{Beta}(1,1) \equiv \text{U}(0,1)$$

If X ~Gamma(α, λ) and Y ~Gamma(β, λ), then $U = X + Y$ ~Gamma$(\alpha + \beta, \lambda)$ and $V = X/(X+Y)$ ~Beta(α, β) and are independent

Chi-Square Distribution

$$\chi^2_n \equiv \text{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right)$$

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

Monotonic Transformation

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| & \text{if } y \in \mathcal{R}(g) \\ 0 & \text{otherwise} \end{cases}$$

2.4 Joint Distribution

$$\begin{aligned}P(a_1 < X \le a_2, b_1 < Y \le b_2) &= F_{X,Y}(a_2, b_2) \\ &- F_{X,Y}(a_1, b_2) + F_{X,Y}(a_1, b_1) - F_{X,Y}(a_2, b_1)\end{aligned}$$

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

$$f_{X,Y}(x, y) = h(x)g(y) \Leftrightarrow X \text{ and } Y \text{ are independent}$$

Convolution of Independent Distributions

$$F_{X+Y}(a) = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y) dy$$

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a - y) f_Y(y) dy$$

Bivariate Normal Distribution

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)$$

Joint Transformation

$$f_{Y_1, Y_2, \ldots, Y_n}(y_1, y_2, \ldots, y_n) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \frac{1}{|J(x_1, x_2, \ldots, x_n)|}$$

where

$$J(x_1, x_2, \ldots, x_n) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix},$$

$Y_i = g_i(X_1, X_2, \ldots, X_n)$, and $x_i = h_i(y_1, y_2, \ldots, y_n)$ for $i = 1, 2, \ldots, n$

2.5 Expectations and Variance

Expectations

- If X and Y are independent, then $E[g(X)h(Y)] = E[g(x)]E[h(y)]$
- $E(X) = E[E[X|Y]]$

Covariance

$$\begin{aligned} Cov(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

$$Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j Cov(X_i, Y_j)$$

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) + 2\sum_{1\le i<j\le n} Cov(X_i, X_j)$$

$$Var(X|Y) = E[X^2|Y] - [E(X|Y)]^2$$

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

Correlation

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

$\rho(X,Y) = \pm 1$ iff $Y = \pm aX + b$ where $a = \frac{\sigma_Y}{\sigma_X}$

Moment Generating Functions

$$M_X(t) = E(e^{tX})$$

$$E(X^n) = \left.\frac{d^n}{dt^n} M_X(t)\right|_{t=0}$$

Taylor Series Expansion

$$g(x) = g(\mu) + \frac{g'(\mu)}{1!}(x-\mu) + \frac{g''(\mu)}{2!}(x-\mu)^2 + \ldots$$

Multiplicative Property If X and Y are independent, then $M_{X+Y}(t) = M_X(t)M_Y(t)$

Uniqueness Property If $\exists t$ such that $M_X(t) = M_Y(t) \forall t \in (-h, h)$ then X and Y have the same distributions

Joint Moment Generating Functions

$$M_{X_1,X_2,\ldots,X_n}(t_1, t_2, \ldots, t_n) = E[e^{t_1X_1 + t_2X_2 + \cdots + t_nX_n}]$$

$$M_{X_i}(t) = E[e^{tX_i}] = M_{X_1,X_2,\ldots,X_n}(0, \cdots, 0, t, 0, \cdots, 0)$$

2.6 Inequalities

$$a \le X \le b \Rightarrow a \le E(X) \le b$$

Monotone Property $X \le Y \Rightarrow E(X) \le E(Y)$

Boole's inequality

$$P\left(\bigcup_{k=1}^{n} A_k\right) \le \sum_{k=1}^{n} P(A_k)$$

Markov's inequality Let X be a nonnegative random variable. For $a > 0$, we have

$$P(X \ge a) \le \frac{E(X)}{a}$$

Chebyshev's inequality For $a > 0$, we have

$$P(|X-\mu| \ge a) \le \frac{Var(X)}{a^2}$$

One-sided Chebyshev's inequality If X is a r.v. with mean 0 and finite variance σ^2, then, for any $a > 0$,

$$P(X \ge a) \le \frac{\sigma^2}{\sigma^2 + a^2}$$

Jensen's inequality If $g(x)$ is a convex function, then

$$E[g(x)] \ge g[E(X)]$$