## - Basic Distributions -

**Binomial Distribution** $X \sim Bin(n, p)$

If x is the number of successes in n independent Bernoulli trials, $(Be(p) = Bin(1, p))$

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$E(X) = np,\ Var(X) = np(1-p),\ M_X(t) = (1 - p + pe^t)^n$

**Geometric Distribution** $X \sim Geo(p)$

Define the random variable $X$ as the number of trials until the first success. Then the probability mass function of $X$ is:

$$P(X = x) = (1-p)^{x-1} p$$

$E(X) = \frac{1}{p},\ Var(X) = \frac{1-p}{p^2},\ M_X(t) = \frac{pe^t}{1-(1-p)e^t}$

**Negative Binomial** $X \sim NB(r, p)$

Define the random variable $X$ as the number of trials until the $r$th success. Then the probability mass function of $X$ is:

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

$E(X) = \frac{r}{p},\ Var(X) = \frac{r(1-p)}{p^2}$

**Hypergeometric** $X \sim HGeom(N, K, n)$

If $X$ is the number of successes in $n$ draws without replacement from a population of $N$ objects of which $K$ are successes, then the probability mass function of $X$ is:

$$P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$E(X) = \frac{nK}{N},\ Var(X) = \frac{nK(N-K)(N-n)}{N^2(N-1)}$

**Poisson** $X \sim Poisson(\lambda)$

If $X$ is the number of events in a fixed interval of time or space ($\lambda$: rate of occurence per unit time), then the probability mass function of $X$ is:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$E(X) = \lambda,\ Var(X) = \lambda,\ M_X(t) = e^{\lambda(e^t - 1)}$$

Poison random variables can be used as approximations for binomial random variables when $n$ is large and $p$ is small enough so that $\lambda = np$ is moderate(usually if $n > 20$ and $np < 15$).

**Uniform** $X \sim U(a, b)$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$E(X) = \frac{a+b}{2},\ Var(X) = \frac{(b-a)^2}{12},\ M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$

**Normal** $X \sim N(\mu, \sigma^2)$

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

For standard normal distribution, $Z \sim N(0, 1)$, we denote its density function $\phi(z)$ and its distribution function $\Phi(z)$. Remember it's imposible to integrate the normal distribution function except the $-\infty \to \infty$ case. The $q$-th quantile of of a random variable $X$ is defined as a number $z_q$ such that $P(X < z_q) = q$.

**Gamma Distribution**

Gamma function is $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, and we say $X$ follows Gamma distribution($X \sim \Gamma(\alpha, \lambda)$) if

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha$, $\lambda > 0$, and $E(X) = \frac{\alpha}{\lambda}$, $Var(X) = \frac{\alpha}{\lambda^2}$.

When $\alpha = 1$, $\Gamma(1, \lambda) = \text{Exp}(\lambda)$, which has the unique property of memorylessness. The Gamma distrubution can also be thought of as a waiting time between Poisson distributed events.

**Beta Distribution**

Beta function is $B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$, and we say $X$ follows Beta distribution($X \sim \text{Beta}(a, b)$) if

$$f_X(x) = \begin{cases} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $a, b > 0$, and $E(X) = \frac{a}{a+b}$, $Var(X) = \frac{ab}{(a+b)^2}$, by showing that $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

**Cauchy** $X \sim Cauchy(\theta)$

$$f_X(x) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2},\ -\infty < x < \infty$$

Cauchy distribution has no mean or variance (both diverges) although it is a valid pdf, and its median is $\theta$. The ratio of two N(0,1) is Cauchy(0).

## - Calculation Tricks -

$f(x) = \sum_{k=0}^\infty \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!} x^2 + \cdots$
$\int xe^{ax} dx = \frac{1}{a^2} e^{ax}(ax - 1)$
$\int x^2 e^{ax} dx = \frac{1}{a^3} e^{ax}(a^2x^2 - 2ax + 2)$
$\int x^3 e^{ax} dx = \frac{1}{a^4} e^{ax}(a^3x^3 - 3a^2x^2 + 6ax - 6)$

# - Transformations of RV -

## Real-valued Functions of RV

For discrete random variables $X$ and $Y$:

$$p_{g(X,Y)}(z) = \sum_{x,y:g(x,y)=z} p_{X,Y}(x,y)$$

For continuous random variables $X$ and $Y$:

$$f_{g(X,Y)}(z) = \frac{d}{dz} \iint_{g(x,y)=z} f_{X,Y}(x,y)\,dx\,dy$$

By calculations, we have these special cases:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)dx$$
$$f_{X-Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, x-z)dx$$
$$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_{X,Y}(x, \frac{z}{x})dx$$
$$f_{X/Y}(z) = \int_{-\infty}^{\infty} |y| f_{X,Y}(zy, y)dy$$

## Vector-valued F of RV (Change of Variables)

Single integral:

$$\int_{x=a}^{x=b} f(x)dx = \int_{y=g(a)}^{y=g(b)} f(g^{-1}(y))\frac{dx}{dy}dy$$

Using the property of monotonicity, when $Y = g(X)$:

$$f_Y(y) = f_X(g^{-1}(y))|\frac{dx}{dy}|$$

Double integral:

$$\iint_R f_(x,y)\,dx\,dy = \iint_S f_(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv$$

where $S$ is the region in the $uv$ plane that corresponds to region $R$ in the $xy$ plane, and

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right|$$
$$= \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1} = \left|\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right|^{-1}$$

Hence, we have

$$f_{U,V}(u,v) = f_{X,Y}(x(u,v), y(u,v))\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$

# - Integration Techniques -

By parts: $\int udv = uv - \int vdu$
Substitution: $\int f(g(x))g'(x)dx = \int f(u)du$

# - Expectation and Variance -

## Expectation

$$E[g(X,Y)] = \begin{cases} \sum_{x,y} g(x,y)p_{X,Y}(x,y) & \text{dis} \sim \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)\,dx\,dy & \text{con} \sim \end{cases}$$

$$E(X) = \sum_x P(X > x) \text{ or } \int_0^{\infty} P(X > x)dx$$

Linearity property (application: Boole's inequality):

$$E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E[X_i]$$

Independence property (converse not true):

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

eg. Coupon-collecting: one type one RV

## Covariance and Correlation

$$\text{cov(X,Y)} = E(X-\mu_X)(Y-\mu_Y) = E(XY) - \mu_X\mu_Y$$
$$\rho(X,Y) = \frac{\text{cov(X,Y)}}{\sigma_X\sigma_Y} \in [-1, 1]$$

Linearity property (No.3 - Independence):

$$\text{cov}(aX+b, cY+d) = ac\cdot\text{cov}(X,Y)$$
$$\text{cov}(\sum_{i=1}^n a_iX_i, \sum_{j=1}^m b_jY_j) = \sum_{i=1}^n\sum_{j=1}^m a_ib_j\text{cov}(X_i,Y_j)$$
$$\text{var}(\sum_{k=1}^n X_k) = \sum_{k=1}^n \text{var}(X_k) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^n \text{cov}(X_i,X_j)$$

## Conditional Expectation and Variance

Expectation:

$$E[X|Y=y] := \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx = \int_{-\infty}^{\infty} x\frac{f_{X,Y}(x,y)}{f_Y(y)}dx$$
$$E[X] = E[E[X|Y]] = \int_{-\infty}^{\infty} E[X|Y=y]f_Y(y)dy$$

Variance:

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

# - Moment Generating Function -

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx}f_X(x)dx$$
$$E(X^n) = M_X^{(n)}(0)\frac{d^n}{dt^n}M_X(t)|_{t=0}$$
$$M_{X,Y}(s,t) = E[e^{sX+tY}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{sx+ty}f_{X,Y}(x,y)dxdy$$

Can be used to determine distribution and independence!