Probability Axioms

1. $P(A) \geq 0$ for $A \in S$ (sample space)
2. $P(S) = 1$
3. If $A_1, A_2 \ldots$ are pairwise disjoint $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Lemmas

a. $P(\emptyset) = 0$

b. $P(A) \leq 1$

c. $P(A^c) = 1 - P(A)$

d. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

e. $A \subset B \Rightarrow P(A) \leq P(B)$

- Bonferroni's Inequality

$P(A \cap B) \geq P(A) + P(B) - 1$ $\quad$ $P(\bigcap_{i=1}^{k} A_i) \geq \sum_{i=1}^{k} P(A_i) - (k-1)$

↳ bounds on intersection of A and B

$$P(A) + P(B) - 1 \leq P(A \cap B) \leq \min(P(A), P(B))$$

- Boole's Inequality

$P(\bigcup A_i) \leq \sum P(A_i)$

$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$ where C_i are a partition of S

Counting
- arrangements of size r from n objects

- without replacement

$\frac{n!}{(n-r)!}$ ordered

$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}$ unordered

- with replacement

n^r ordered

$\binom{n+r-1}{r}$ unordered

- Multinomial n objects, m types repeated $r_1, r_2, \ldots r_m$ times

$\frac{n!}{r_1!\, r_2! \cdots r_m!}$ ordered samples

Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A, B)}{P(B)}$$

Independence
A and B are independent if

$P(A, B) = P(A)P(B)$

Mutual Independence A_i, $i = 1, \ldots, k$ are mutually independent if

$P(\bigcap_{i \in J} A_i) = \prod_{i \in J} A_i$ for any subset J of A_i

Bayes Rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

w/ conditional prob

$$P(A|B,C) = \frac{P(B|A,C)P(A,C)}{P(B|A,C)P(A|C) + P(B|A^c,C)P(A^c|C)}$$

Random variable - function from sample space to the real line

$$P(X=x) = P(\{s \in S : X(s) = x\})$$

probability mass function (pmf) - probability for event x in discrete r.v.

$P(X=x) = f_X(x)$ is a pmf if:

a) $f_X(x) \geq 0 \ \forall x$

b) $\sum_x f_X(x) = 1$

probability distribution function (pdf) - function of continuous r.v.

$\int_{-\infty}^{x} f_X(u)\,du = F_X(x)$. f_X is pdf if:

a) $f_X(x) \geq 0 \ \forall x$

b) $\int_{-\infty}^{\infty} f_X(x)dx = 1$

cumulative distribution function (cdf) - sum of probabilities less than or equal to x

$P(X \leq x) = F_X(x)$ is cdf if:

a) $\lim_{x \to -\infty} F(x) = 0 \quad \lim_{x \to \infty} F(x) = 1$

b) $F(x)$ is non-decreasing

c) $F(x)$ is right continuous

$$\lim_{x \to x_0^+} F(x) = F(x_0)$$

Geometric series

$$S_n = \sum_{i=1}^{n} q^{i-1} = 1 + q + q^2 + \cdots$$
$$= \frac{1-q^n}{1-q}$$

for continuous r.v.

$\frac{d}{dt} F_X(t) = f_X(t) \qquad f_X(x) \neq P(X=x)$

$$P(a \leq X \leq b) = \int_a^b f_X(t)\,dt = F_X(b) - F_X(a)$$

X + Y are identically distributed if

$F_X(w) = F_Y(w)$ does not imply $X = Y$

Transformations of random variables

$X \sim F_X(x)$ $\quad Y = g(X)$ where $g(\cdot)$ is monotone and X continuous

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

If $g(\cdot)$ non-monotone, define monotone sections, apply formula for each and add

for cdf use

$F_Y(y) = P(Y \le y) = P(g(X) \le y)$ solve for prob statement involving X

Probability Integral Transformation

$X \sim F_X(x)$ $\quad X$ continuous

$Y = g(X) = F_X(X)$ plug X into it's own cdf

$Y \sim Unif(0,1)$

Inverse cdf (quantile function)

set cdf equal to q, solve for x $[F(x) = q]$

$F_X^{-1}(q) = \inf\{x : F_X(x) \ge q\}$ $\quad 0 < q < 1$

Expected Values

$$E[X] \begin{cases} \int_{-\infty}^{\infty} x f_X(x)\, dx & \text{continuous} \\ \sum_x x f_X(x) & \text{discrete} \end{cases}$$

often useful to find kernel of known density and integrate or sum over support to reduce to 1

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

Expectation properties

$E[aX + b] = aE[X] + b$

$E[(X-b)^2]$ is minimized if $b = E[X]$

Moments

k^{th} moment $E[X^k]$

k^{th} central moment $E[(X - E[X])^k]$

2^{nd} central moment $Var(X) = E[(X-\mu)^2]$
$= E[X^2] - (E[X])^2$

$Var(aX+b) = a^2 Var(X)$

Moment Generating Function (mgf)

$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$ $\qquad \frac{d^k}{dt^k} M_X(t) \Big|_{t=0} = E[X^k]$

mgf properties

If $M_X(t) = M_Y(t)$ then $F_X(x) = F_Y(x) \quad \forall x$

If $E[X^k] = E[Y^k] \ \forall k$ then $F_X(x) = F_Y(x) \ \forall x$ if $X, + Y$ have bounded support

Discrete Distributions

- Bernoulli $f(x) = p^x(1-p)^{1-x}$ $x=0,1$ $0<p<1$ $E[X]=p$ $Var[X] = p(1-p)$
- Binomial $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ $x=0,1,\dots n$ $0<p<1$ $E[X]=np$ $Var[X]=np(1-p)$
- Geometric $f(y) = (1-p)^{y-1}p$ $y=1,2,\dots$ $0<p<1$ $E[Y]=\frac{1}{p}$ $Var[Y] = \frac{1-p}{p^2}$
- Negative Binomial $f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$ $x=r, r+1, \dots$ $0<p<1$ $E[X]=\frac{r}{p}$ $Var[X] = \frac{r(1-p)}{p^2}$
  ↳ x trials for r successes
- Discrete Uniform $f(x) = \frac{1}{N}$ $x = 1, 2, \dots, N$ $E[X] = \frac{N+1}{2}$ $Var[X] = \frac{(N+1)(N-1)}{12}$
- Poisson $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $x=0,1,2,\dots$ $\lambda \geq 0$ $E[X]=\lambda$ $Var[X]=\lambda$
- Multinomial
  m trials
  cell probs p_i
  $f(x_1, \dots x_n) = \frac{m!}{x_1! \cdots x_n!} p_1^{x_1} \cdots p_n^{x_n}$ $\quad \Sigma x_i = m \quad \Sigma p_i = 1$

  $\Rightarrow$ Each X_i (marginal) $\sim Bin(m, p_i)$ $\qquad Cov(X_i, X_j) = m p_i p_j \quad i \neq j$

Continuous Distributions

- Uniform $f(x) = \frac{1}{b-a}$ $a<X<b$ $E[X] = \frac{a+b}{2}$ $Var[X] = \frac{(b-a)^2}{12}$
- Gamma $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$ $x>0$ $\alpha, \beta > 0$ $E[X] = \alpha\beta$ $Var[X] = \alpha\beta^2$
- Chi-squared
  p degrees of freedom $f(x) = \frac{1}{\Gamma(\frac{p}{2}) 2^{p/2}} x^{p/2 - 1} e^{-x/2}$ $E[X] = p$ $Var[X] = 2p$
  ↳ Gamma$(\alpha = \frac{p}{2}, \beta = 2)$
  if $Z \sim Normal(0,1)$ $X = Z^2 \sim \chi_1^2$
- Exponential $f(x) = \frac{1}{\beta} \exp[-x/\beta]$ $x>0$, $\beta > 0$ $E[X] = \beta$ $Var[X] = \beta^2$
  ↳ Gamma$(\alpha = 1, \beta)$
  memory-less property $P(X>t \mid X>s) = P(X>t-s)$ for $0<s<t$

Continuous distributions (cont.)

- Normal $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$ $-\infty < x < \infty$, $\sigma > 0$, $-\infty < \mu < \infty$ $E[X] = \mu$ $Var[X] = \sigma^2$

- Beta $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ $0 < x < 1$ $\alpha, \beta > 0$ $E[X] = \frac{\alpha}{\alpha+\beta}$
  $Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

- Cauchy $f(x) = \frac{1}{\pi}\left(\frac{1}{1+(x-\theta)^2}\right)$ $F^{-1}(1/2) = \theta$ $E[X] = \infty$ (does not exist)

__Exponential family__ probability function can be written in form

$$f(x|\theta) = h(x)\, c(\theta) \exp\left[\sum_{i=1}^{k} w_i(\theta)\, t_i(x)\right]$$

$h(x), t_i(x)$ cannot depend on θ $c(\theta), w_i(\theta)$ cannot depend on X

→ Any distribution whose support depends on its parameters is __not__ in exponential family

__Location-scale family__

$X \sim f_X(x)$ $Y = \sigma X + \mu$ $\sigma > 0$, $-\infty < \mu < \infty$

$$f_Y(y) = \frac{1}{\sigma} f_X\left(\frac{y-\mu}{\sigma}\right)$$

__Joint + Marginal Distributions__

- discrete joint pmf is
  $f(x,y) = P(X=x, Y=y)$
  $f_X(x) = P(X=x) = \sum_y f(x,y)$
  $f_Y(y) = P(Y=y) = \sum_x f(x,y)$

- continuous $f(x,y)$ is joint pdf if
  $P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$ $\forall A$
  $f_X(x) = \int_x f(x,y)\,dx$
  $f_Y(y) = \int_y f(x,y)\,dy$

- marginal pdfs do not define a unique joint pdf

$$E[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y) f(x,y) & \text{discrete} \\ \int_y \int_x g(x,y) f(x,y)\,dx\,dy & \text{continuous} \end{cases}$$

__Conditional Distributions__

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x)\,dx \quad \text{or} \quad \sum x f_{X|Y}(x)$$

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

Independent Distributions X and Y are independent if for x and $y \in \mathbb{R}$

$$f(x,y) = f_X(x) f_Y(y)$$

X and Y are independent iff $\exists\, g(x), h(y)$ s.t for $x \in \mathbb{R}, y \in \mathbb{R}$

$$f(x,y) = g(x)\,h(y)$$

if $X \perp\!\!\!\perp Y$

a) $E[g(X)h(Y)] = E[g(X)]\,E[h(Y)]$

b) $M_X(t) M_Y(t) = M_Z(t)$

where $M_Z(t)$ is the mgf of $Z = X+Y$

Bivariate Transformations let (X,Y) be bivariate r.v. and consider new r.v. (U,V) such that $U = g_1(X,Y)$ $\quad V = g_2(X,Y)$ $\Rightarrow X = g_1^{-1}(U,V)$ $\quad Y = g_2^{-1}(U,V)$

discrete

$$f_{U,V}(u,v) = f_{X,Y}(g_1^{-1}(U,V), g_2^{-1}(U,V))$$

continuous

$$f_{U,V}(u,v) = f_{X,Y}(g_1^{-1}(U,V), g_2^{-1}(U,V))\,|J| \quad \text{where } J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{dx}{du}\frac{dy}{dv} - \frac{dy}{du}\frac{dx}{dv}$$

where $u = g_1(X,Y)$ $\quad v = g_2(X,Y)$ is a one-to-one, onto transfunction, o.w. find partition where one-to-one + sum

Hierarchical Models + Mixture Distributions

a r.v. X has a mixture distribution if X depends on a quantity that also has a distribution

Ex. $X|Y \sim Bin(Y,p)$ $\quad Y \sim Pois(\lambda)$

$$f(x) = \int f(x,y)\,dy = \int f(x|y) f(y)\,dy$$

$$E[X] = E[E(X|Y)] \qquad E[g(X)] = E[E(g(X)|Y)]$$

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

Covariance and Correlation for r.v. $X + Y$

$$Cov(X,Y) = E[(X - E(X))(Y - E[Y])] = \sigma_{XY}$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \rho_{XY}$$

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

Covariance properties

a) $Cov(aX+bY, cW+dZ) = ac\,Cov(X,W) + ad\,Cov(X,Z) + bc\,Cov(Y,W) + bd\,Cov(Y,Z)$

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

c) $Cov(X,X) = Var(X)$

d) $Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\,Cov(X,Y)$

e) $$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{i\neq j}\sum a_i a_j Cov(X_i, X_j)$$
$$= \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} a_i a_j Cov(X_i, X_j)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$$

f) If $X \perp\!\!\!\perp Y$, then $Cov(X,Y) = 0$

NB: $Cov(X,Y) = 0 \not\Rightarrow X \perp\!\!\!\perp Y$

g) $|\rho_{XY}| = 1$ iff $\exists\, a \neq 0, b$ s.t. $P(Y = aX + b) = 1$

Bivariate Normal

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \boldsymbol{\Sigma}\right) \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix} \qquad \rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \Rightarrow \sigma_{XY} = \rho\sigma_X\sigma_Y$$

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}\right]$$

$f(x) = \int f(x,y)\,dy \sim N(\mu_X, \sigma_X^2)$

$f(y) = \int f(x,y)\,dx \sim N(\mu_Y, \sigma_Y^2)$

$f(y|x) \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \sigma_Y^2(1-\rho^2)\right)$

Multivariate Distributions

$f(\underline{x}) = f(x_1, \dots x_n)$

conditional

$f(x_1) = \int\cdots\int f(\underline{x})\,dx_2\cdots dx_n$

$f(x_2,\dots,x_n) = \int f(\underline{x})\,dx_1$

marginal

$F(x_1|x_2,\dots,x_n) = \dfrac{f(\underline{x})}{f(x_2,\dots,x_n)}$

$X_1,\dots,X_n$ are mutually independent if $f(x_1,\dots,x_n) = f(x_1)\cdots f(x_n) = \prod_{i=1}^{n} f(x_i)$

If $X_1,\dots,X_n$ are mutually ind. w/ mgfs $M_{X_1}(t),\dots,M_{X_n}(t)$ and $Z = X_1 + \cdots + X_n$

$M_Z(t) = M_{X_1}(t)\cdots M_{X_n}(t)$ if X_i are identically distributed $M_Z(t) = (M_X(t))^n$

Chebyshev's Inequality

$$P(g(x) \geq r) \leq \frac{E[g(x)]}{r}$$

Applied form

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

Cauchy-Schwartz Inequality

$$|E[XY]| \leq E[|XY|] \leq (E[|X|^2]^{1/2})(E[|Y|^2])^{1/2}$$

Jensen's Inequality

$E[g(X)] \geq g(E[X])$ if $g(X)$ is convex ($g''(x) \geq 0 \ \forall\ x$)

Direction of inequality reverses if $g(X)$ is concave

Equality holds if $g(X)$ is a line or a point mass at $E[X]$

Random Samples

random variables are independent + identically distributed (iid) if all are mutually independent and have same distribution

A sample from an infinite population or a sample from a finite population with replacement is iid.

A sample from a finite population without replacement is not iid.

Let $X_1, \ldots, X_n$ be a sample from a population and let

$$Y = T(X_1, \ldots, X_n) \text{ for function } T(\cdot)$$

Y is a statistic and the probability distribution of Y is the sampling distribution

sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}(\Sigma x_i^2 - n\bar{x}^2)$

Let $X_1, \ldots, X_n$ be a random sample from a population + $g(X)$ be a fun. such that $E[g(X_1)]$ and $Var[g(X_1)]$ exist then:

$$E[\Sigma g(X_i)] = n E[g(X_1)] \qquad Var(\Sigma g(X_i)) = n\, Var(g(X_1))$$

Let pop. mean $= \mu$ and pop. variance $= \sigma^2 < \infty$ for random sample $X_1, \ldots, X_n$

$$E[\bar{X}] = \mu \qquad Var[\bar{X}] = \frac{\sigma^2}{n} \qquad E[S^2] = \sigma^2$$

- Let $X_1, \ldots, X_n$ be a random sample from a population w/ mgf $M_X(t)$ then the mgf of the sample mean is

$$M_{\bar{X}}(t) = [M_X(t/n)]^n$$

- Suppose $X_1, \ldots, X_n$ is a random sample from a pdf or pmf that is a member of an exponential family so

$$f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right)$$

statistics $T_1, \ldots, T_k$ where $T_i(X_1, \ldots, X_n) = \sum_{j=1}^{n} t_i(X_j)$ $i = 1, \ldots, k$ is an exponential family if $\{w_1(\theta), w_2(\theta) \ldots w_k(\theta), \theta \in \Theta\}$ contains an open subset of $\mathbb{R}^k$ where

$$f_T(u_1, \ldots, u_k|\theta) = H(u_1, \ldots u_k)[c(\theta)]^n \exp\left[\sum_{i=1}^{k} w_i(\theta)u_i\right]$$

Normal Distribution properties for $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$:

1. $\bar{X} \sim N(\mu, \sigma^2/n)$
2. $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$
3. $\bar{X} \perp\!\!\!\perp S^2$

Chi-squared properties

1. if $Z \sim N(0,1)$ then $Z^2 \sim \chi^2_1$
2. if $X_i \overset{ind}{\sim} \chi^2_{p_i}$ then $\sum X_i \sim \chi^2_{\sum p_i}$

t-distribution

for $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ $\quad \dfrac{\bar{X} - \mu}{S/\sqrt{n}} = \dfrac{(\bar{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{S^2/\sigma^2}} \sim t_{n-1}$

form is $U/\sqrt{V/p}$ where $U \sim N(0,1)$ $V \sim \chi^2_p$ $U \perp\!\!\!\perp V$

F-distribution Let $X_i \overset{iid}{\sim} N(\mu_X, \sigma_X^2)$ $i = 1, \ldots, n$ $Y_j \overset{iid}{\sim} N(\mu_Y, \sigma_Y^2)$ $j = 1, \ldots, m$

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F_{n-1, m-1}$$

form is $(U/p)/(V/q)$ where $U \sim \chi^2_p$ $V \sim \chi^2_q$ $U \perp\!\!\!\perp V$

Order Statistics $X_i \overset{iid}{\sim} F(x)\ i=1,\dots,n$

$X_{(1)}$ = smallest X_i = $\min[X_i]$ "sample min"

$X_{(n)}$ = largest X_i = $\max[X_i]$ "sample max"

cdf of k^{th} order stat $X_{(k)}$

$$F_{X_{(k)}}(x) = \sum_{i=k}^{n} \binom{n}{i} [F(x)]^i [1-F(x)]^{n-i}$$

pmf of k^{th} order stat (discrete)

$$f_{X_{(k)}}(x_j) = P(X_{(k)} = x_j) = F_{X_{(k)}}(x_j) - F_{X_{(k)}}(x_{j-1}) =$$
$$\sum_{i=k}^{n} \binom{n}{i} \left[F(x_j)^i (1-F(x_j))^{n-i} - F(x_{j-1})^i (1-F(x_{j-1}))^{n-i}\right]$$

pdf of k^{th} order stat (continuous)

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} f(x) (1-F(x))^{n-k}$$

pdf joint dist. of $X_{(k)}, X_{(\ell)}$ (continuous)

$$f_{X_{(k)}, X_{(\ell)}}(y_1, y_2) = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} F(y_1)^{k-1} f(y_1) [F(y_2) - F(y_1)]^{\ell-k-1} f(y_2) [1-F(y_2)]^{n-\ell}$$

joint pdf of all order statistics

$$f_{X_{(1)} \cdots X_{(n)}}(y_1, \dots y_n) = n!\, f(y_1) f(y_2) \cdots f(y_n) \quad \text{for } -\infty < y_1 < y_2 < \cdots y_n < \infty$$

Convergence in Probability

for sequence $X_1, X_2, X_3, \dots$ with $\varepsilon > 0$

$$X_n \xrightarrow{P} X \iff \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{or} \quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

Weak law of large numbers

if $X_i \overset{iid}{\sim} G(\mu, \sigma^2)$ then $\bar{X}_n \xrightarrow{P} \mu$

Continuous mapping Theorem

If $X_n \xrightarrow{P} X$ and $h(\cdot)$ is a continuous function then $h(X_n) \xrightarrow{P} h(X)$

Almost sure convergence

$$X_n \xrightarrow{a.s.} X \iff P(\lim_{n\to\infty} |X_n - X| \geq \varepsilon) = 0 \quad \text{or} \quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

strong law of large numbers

$$\bar{X}_n \xrightarrow{a.s.} \mu$$

Convergence in Distribution

$X_n \xrightarrow{d} X \iff \lim_{n\to\infty} F_{X_n}(x) = F_X(x) \Rightarrow \lim_{n\to\infty} P(X_n \le x) = P(X \le x)$

Convergence relationships

$X_n \xrightarrow{as} X \Rightarrow X_n \xrightarrow{P} X$

$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

Central Limit Theorem

Let $X_1, X_2, \ldots$ be sequence of iid r.v. $E[X_i] = \mu$ $Var(X_i) = \sigma^2 > 0$

$\bar{X}_n = \frac{1}{n}\sum X_i$ with $G_n(x)$ the cdf of $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ then for any x

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy \quad \text{i.e.} \quad \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

Slutsky's Theorem

if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$ then

$X_n Y_n \xrightarrow{d} cX$

$X_n + Y_n \xrightarrow{d} X + c$

Delta Method

If $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ and there is $g(X)$ where $g'(X)$ exists and $g'(\theta) \neq 0$ then

$$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2 \sigma^2)$$

Second order Delta Method

$g'(\theta) = 0$ and $g''(\theta)$ exists and is not 0

$$n[g(Y_n) - g(\theta)] \xrightarrow{d} \sigma^2 \frac{g''(\theta)}{2} \chi^2_1$$

Taylor series expansion of g around θ

$$g(t) = \underbrace{\underbrace{g(\theta) + g'(\theta)(t-\theta)}_{\text{first order}} + \frac{g''(\theta)}{2}(t-\theta)^2}_{\text{second order}} + \text{remainder}$$