Probability

- **Sample Space**: The set of all possible outcomes of an experiment. eg. S = {HH, HT, TH, TT}
- **Equally likely outcome**: If there are m equally likely possibilities, of which one must occur and s are regarded as a success, then the prob of a success is given by s=m.
- **Frequency Interpretation**: The prob of an event is the proportion of times the event will occur in a long run of repeated experiments.
- **Personal Prob**: the degree to which a given individual believes that the event will happen. Sometimes, the term subjective prob is used because the degree of belief may be different for each individual.
- **Union and Intersection**: As per normal set union and intersection definitions.
  Union and Intersection are commutative, associative and distributive. Think of intersection as multiplication and union as addition.
  De Morgan's laws also apply.
  $(E_1 \cup E_2 \cup ... E_n)^c = E_1^c \cap E_2^c \cap ... E_n^c$, vice versa.
- **Generalised Basic Principle of Counting: Multiplication principle**: Suppose that r tests are performed, each with n_i choices. Then there are $\prod_{i=0}^{r} n_i$ combinations to do all of the tests.
  For example: 10 MCQ qns with 4 choices each. Then 3^{10} ways of getting all answers wrong.
- **Factorial**: If there are n distinct objects, the number of ways of arranging them is n!
- **Circles**: For n distinct items in a circle, there are $\frac{n!}{n}$ ways to arrange them.
- **Necklaces**: There are $(n-1)!/2$ ways to arrange items on a necklace.
- **Generalised Basic Principle of Counting: Addition Principle**: Suppose that r tests are performed, each with n_i choices. Then there are $\sum_{i=0}^{r}$ ways to perform any 1 experiment (a or b or c...).
- **Permutations: distinct objects**: Number of ways to arrange r out of n distinct objects:
  $nPr = \frac{n!}{(n-r)!}$. Order matters.
- **Permutations: not all distinct**: Let there be n_i copies of the ith object. Number of ways to arrange these n objects of k types: $\frac{n!}{n_1!n_2!...n_k}$
- **Combinations: distinct objects**: Number of ways to choose r objects from n distinct ones: nCr $= \frac{n!}{r!(n-r)!}$
- **Prob**:
  1. For any event A, $0 \leq P(A) \leq 1$.
  2. Let S be the sample space. $P(S) = 1$.
  3. For any mutually exclusive events $A_1, A_2, ...$, the union of their probabilities = sum of probabilities, while intersection = 0.
  4. $P(A^c) = 1 - P(A)$
  5. If $A \subseteq B$, $P(A) + P(BA^c) = P(B)$
  6. (Incl-excl principle):
  $P(A \cup B) = P(A) + P(B) - P(AB)$. 7.
  $P(A \cup B \cup C) = P(A) + P(B) + P(C) - [P(AB) + P(AC) + P(BC)] + P(ABC)$ 8.
  $P(A) = \frac{\text{number of ways in A}}{\text{number of ways in S}}$
- **Conditional Prob**: B given that A has occurred: $P(B|A) = \frac{P(B \cap A)}{P(A)}$
  Remark: $P(\cdot|A)$ is also a prob, so all properties of prob apply too.

- **Multiplication Rule**:
  $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$
- **Generalised Multiplication Rule**:
  $P(A_1 A_2 ... A_n) =$
  $P(A_1)P(A_2|A_1)P(A_3|A_1A_2)...P(A_n|A_1A_2...A_{n-1})$
- **Inverse Prob**: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$
- **Indep**: Two events A and B are indep if and only if $P(A \cap B) = P(A)P(B)$ or $P(A|B) = P(A)$ or $P(B|A) = P(B)$.
- **Mutual Exclusion**: Two events A and B are mutually exclusive if and only if $P(A \cap B) = 0$.
- **Theorem on independence**: If A and B are indep, so are A and B^c, A^c and B, A^c and B^c.
- **Pairwise Indep**: A set of events $A_1, A_2, ..., A_n$ are pairwise indep if and only if $P(A_i A_j) = P(A_i)P(A_j)$ for all i,j, $i \neq j$.
- **Mutual Indep**: A set of events $A_1, A_2, ..., A_n$ are mutually indep if and only if for any subset k, all the events are indep, ie
  $P(A_{i\ ink}) = P(A_i)...P(A_k)$. There are $2^n - n - 1$ cases.
- **Remarks**:
  1. Mutual independence implies pairwise independence, but not vice versa. 2. Suppose $A_1, ...A_n$ are mutually indep events, and $B_i = A_i or A_i^c$ for $i = 1, ..., n$. Then $B_1, ...B_n$ are also mutually indep.
- **Bayes' Theorem**: Let $B_1, ..., B_n$ be a **partition** of S. For any event A, and any $k \in 1, ..., n$

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{P(B_1)P(A|B_1) + ... + P(B_n)P(A|B_n)}$$

- **Example**: Calculate $P(A^c|O)$

$$P(A) = 3/4, \quad P(O|A) = 17/20, \quad P(O|A^c) = 3/20.$$

The required probability will be given as

$$\begin{aligned} P(A^c|O) &= \frac{P(O|A^c)P(A^c)}{P(O|A^c)P(A^c) + P(O|A)P(A)} \\ &= \frac{3/20 \times 1/4}{3/20 \times 1/4 + 17/20 \times 3/4} \\ &= 3/54 = 1/18 \\ &\approx 0.05556. \end{aligned}$$

- **Rule of Total Prob**:
  $P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + ...$
  where $B_1, B_2, ...$ is a partition of S.

Random variables

- **Random Variable**: A random variable X is a mapping from the sample space S to the set of real numbers $\mathbb{R}$. That is: $X : S \to \mathbb{R}$.
  We use uppercase for the variable: X, and lowercase for a value x that the variable can take.
- **Equivalent Event**: Let $A = \{s \in S | X(s) \in B\}$.
  In other words, A consists of all sample points in which X(s) is in B. A is the pre-image of B.
  $P(A) = P(B)$, so we say A and B are equivalent events.
- **Prob Mass Function for Discrete Variables**: Prob function $f_X(x_i) = P(X = x_i)$.
- **Properties of PMF for Discrete Variables**:
  1. $f(x_i) \geq 0$ for every x_i
  2. $\sum_{x_i} f(x_i) = 1$
  3. $P(X \in E) = \sum_{x_i \in E} f(x_i)$
- **Example**: Let $p(k) = c\frac{\lambda^k}{k!}, k = 0, 1, 2, ...$
  Then $\sum_{k=0}^{\infty} p(k) = 1$.
  Thus $1 = c\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = ce^{\lambda}$ (by Maclaurin's Series).
  Thus $c = e^{-\lambda}$
- **Prob Function**: Can be expressed in multiple ways.
  1. $f(x) = \begin{cases} kx^2, & \text{if } x \geq 1 \\ 0, & \text{otherwise} \end{cases}$
  2.

y	-1	1	2	3
f(y)	0	1	4	9

- **Prob Mass Function for Cont Variable**: $P(a < X \leq b) = \int_a^b f(x)dx$.
- **Properties of PMF for Cont Variable**:
  1. $P(X = x) = 0$ for any x.
  2. $f(x) \geq 0$ for any x.
  3. The total area under the curve is 1, i.e. $\int_{-\infty}^{\infty} f(x)dx = 1$
- **Example**:
  Find k such that f(x) is prob density function.

$$f(x) = \begin{cases} kx^2(1-x), & \text{for } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Solution: Ensure that f(x) integrates to 1.

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_0^1 kx^2(1-x)dx$$

$$= k[\frac{x^3}{3} - \frac{x^4}{4}]_0^1 = k/12$$

For example, the prob that $x \leq 0.5$ is $\int_{-\infty}^{0.5} f(x)dx$

- **Cumulative Distribution Function**: $F(x) = P(X \leq x)$. We usually denote CDF with caps F, and PDF with lowercase f.
- **Properties of CDF**:
  1. $F(x)$ is a non-decreasing function.
  2. $\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1$
  3. $0 \leq F(x) \leq 1$.
- **Discrete CDF**: $\sum_{t \leq x} f(t) = \sum_{t \leq x} P(X = t)$
  Useful formula: $P(a \leq X \leq b)$
  $= P(X \leq b) - P(X < a) = F(b) - F(a^-)$

EXAMPLE 2.23 (GEOMETRIC CDF)
For $0 < p < 1$, let the probability mass function of X be given by

$$f(x) = \begin{cases} (1-p)^{x-1}p, & x = 1, 2, 3, \dots \\ 0, & \text{otherwise} \end{cases}.$$

Find its cumulative distribution function .

Solution:
For any positive integer x, its cumulative distribution function is

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \sum_{k=1}^{x} P(X = k) \\ &= \sum_{k=1}^{x} (1-p)^{k-1}p = p \times \frac{1-(1-p)^x}{1-(1-p)} \\ &= 1-(1-p)^x, \text{ for } x = 1, 2, 3, \dots. \end{aligned}$$

Note that we are not done yet, since $F(x)$ is defined for all $x \in \mathbb{R}$.

As $f(x) = 0$ between positive integers, $F(x)$ is constant between positive integers. Thus

$$F(x) = \begin{cases} 1-(1-p)^{[x]}, & x \geq 1 \\ 0, & x < 1 \end{cases},$$

where $[x]$ is the largest integer $\leq x$. That is $[2.7] = 2$, $[3] = 3$.

- **Cont CDF**: $F(x) = \int_{-\infty}^{x} f(t)dt$.
  If the derivative exists, $f(x) = \frac{d}{dx}F(x)$

Mean and Variance

- **Mean** μ_X: Also known as expected value E(X). Is the weighted average.
- **Discrete Mean**:
  $\mu_X = E(X) = \sum_{x_i} x_i P(X = X_i) = \sum_x x f(x)$.
  In other words, add across all x: $P(x) * val(x)$.
- **Discrete Mean**: $\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x)dx$.
  In other words, integrate area, multiplied by val(x).
- **Properties of expectation**:
  1. If a and b are constants, $E(a + bx) = a + bE(x)$.
- **Expectation of g(x) (IMPT!)**: Given prob density function $f_X(x)$ and any other function g(X) on this random variable X,
  1. Discrete X: $E[g(X)] = \sum_x g(x)f_X(x)$
  2. Cont X: $E[g(x)] = \int_{-\infty}^{\infty} g(x)f_X(x)$
  In other words, just replace the original x with $g(x)$. Don't need to modify $f_X(x)$.
- **k-th moment of x**: Set $g(x) = x^k$.
  In other words, $E(X^k)$ is the k-th moment of X.
- **Variance**: As above, set $g(x) = (x - \mu_x)^2$.
  In other words, $V(X) = \sigma^2 = E[(X - \mu_x)^2]$.
- **Properties of variance**:
  1. $V(X) \geq 0$
  2. (impt) $V(X) = E(X^2) - [E(X)]^2$
  3. If $V(X) = 0$ then $P(X = \mu_X) = 1$.
  4. For constants a, b, $V(a + bX) = b^2 V(X)$
- **Standard Deviation**: $\sigma_X = SD(X) = \sqrt{V(X)}$
- **Chebyshev's Inequality**: Gives an **upper bound** on prob of getting a value that deviates from μ by a certain amount $k\sigma$. Formally,

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

Remark: applying k = 2, there's at most 25% chance that a value is outside 2 S.D. from mean.

2D Random Variables

- **Joint Prob (Density) Fn**: $f_{X,Y}(x, y)$ represents $Pr(X = x, Y = y)$. $f_{X,Y}(x, y) \geq 0$.
  Discrete: $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = 1$
  Cont: $\int\int_{(x,y) \in R_{X,Y}} f_{X,Y}(x, y)\, dx\, dy = 1$ or
  $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$
- **Marginal Dist**: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy$
- **Conditional Dist of Y given $X = x$**:
  $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$ if $f_X(x) > 0$. i.e. fix one dimension and calculate the 1D prob density fn of the other. Note: conditional dist also satisfies normal prob rules e.g. $\int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = 1$
- **X and Y are indep if and only if**:
  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all x, y, or equivalently, $f_{X|Y}(x|y) = f_X(x)$.
- **Expectation of $g(x, y)$**:
  Discrete: $E[g(X, Y)] = \sum_x \sum_y g(x, y)f_{X,Y}(x, y)$
  Cont: $E[g(X, Y)] = \int\int g(x, y)f_{X,Y}(x, y)\, dy\, dx$
- **Linearity of Expectaion**: For **any** $x_1, x_2, ...$,
  $E(a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n) =$
  $a_0 + a_1E(x_1) + a_2E(x_2) + ... + a_nE(x_n)$
- **Covariance**:
  $Cov(X, Y) = \sigma_{X,Y} = E[(X - \mu_x)(Y - \mu_y)]$
- **Properties of Covariance**:
  1. $Cov(X, Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_x\mu_y$
  2. $Cov(X, X) = V(X)$
  3. $Cov(X, Y) = Cov(Y, X)$
  4. $Cov(aX + b, cY + d) = acCov(X, Y)$
  5. $V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCov(X, Y)$
  6. X,Y indep $\implies$ $Cov(Y, Y) = 0$.

- **Correlation Coefficient**:
$Cor(X,Y)=\rho_{X,Y}=\frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$
$-1\leq\rho_{X,Y}\leq 1$.
$X, Y indep \implies \rho_{X,Y}=0$.

Prob Distributions

- **Uniform Discrete Distribution**:
PMF: $f_X(x)=P(X=x)=1/k\ \forall\ x=x_1,x_2,...x_k$
$\mu=E(X)=1/k\sum_{i=1}^{k}x_i$
$\sigma^2=V(X)=1/k\sum_{i=1}^{k}(x_i-\mu)^2$
$\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$, $\sum_{k=1}^{n}k^3=\frac{n^2(n+1)^2}{4}$
- **Bernouli Distribution**: A random variable X has a Bernouli Distribution with parameter $0<p\leq 1$ when it has prob mass function
$f_X(x)=p^x(1-p)^{1-x}$ for $x=0,1$.
$E(X)=p$, $V(X)=p(1-p)$
N. B. Bernouli distribution is a special case of Binomial with n = 1.
- **Binomial Dist** $X\sim B(n,p)$: A random variable X has a Binomial Distribution with parameters $n\in\mathbb{Z}^+$ and $0<p\leq 1$ when it has PMF
$f_X(x)=\binom{n}{x}p^x(1-p)^{1-x}$ for $x=0,1,...,n$
"Make n tries with prob p. X = How many successes?"
$E(X)=np$, $V(X)=np(1-p)$
Note:
1. Trials must be indep.
2. P(success) must be a constant.
3. There can only be success OR failures.
- **Geometric Dist** $X\sim Geom(p)$: A random variable X has a Geometric Distribution with parameters $0<p\leq 1$ when it has PMF
$f_X(x)=(1-p)^{x-1}p$ for $x=1,2,...,n$
"Keep trying (with success rate p) until 1st success. What's the prob that X tries are needed?"
$E(X)=\frac{1}{p}$, $V(X)=\frac{1-p}{p^2}$, $P(X\leq x)=1-(1-p)^x$
Memoryless: $P(X>n+k|X>n)=P(X>k)$.
- **Negative Binomial Dist** $X\sim NB(k,p)$: A random variable X has a Negative Binomial Distribution with parameters $k\in\mathbb{Z}^+$ and $0<p\leq 1$ when it has PMF
$f_X(x)=\binom{x-1}{k-1}p^k(1-p)^{x-k}$ for $x=0,1,...,n$
"Keep trying (with success rate p) until k successes. What's the prob that x tries are needed?"
$E(x)=\frac{k}{p}$, $V(X)=\frac{(1-p)(k)}{p^2}$
Geometric distribution is a special case of NB distribution with $k=1$, i.e. $Geom(p)=NB(1,p)$
- **Poisson Dist**: The PMF of the number of successes X in a Poisson experiment with parameter $\lambda>0$ denoted by $X\tilde{\ }Poisson(\lambda)$ is
$f_X(x)=\frac{e^{-\lambda}\lambda^x}{x!}$ for $x=0,1,2,...$
$E(X)=\lambda$, $V(X)=\lambda$
Here, $E(X)=\lambda$ is the average number of successes occurring in the given time interval or specified region.
- **Binomial → Poisson**: Let $X\sim B(n,p)$. X will have an approx Poisson dist with parameter np.

$$\lim_{n\to\infty,p\to\infty}P(X=x)=\frac{e^{-np}(np)^x}{x!}$$

Approximation good when
$n\geq 20\ and\ p\leq 0.05\ OR\ n\geq 100\ and\ np\leq 10$

If p is large, we can swap p with $1-p$
- **Cont Uniform Dist**: X follows a random uniform distribution with range $[a,b]$, denoted by $X\sim U(a,b)$ if its prob density function is:
$f_X(x)=\frac{1}{b-a}$ for $a\leq x\leq b$.
E(X) $=\frac{a+b}{2}$, V(X) $=\frac{(b-a)^2}{12}$.
- **Exponential Dist**: X follows a exponential distribution with parameter λ denoted by $X\sim Exp(\lambda)$ if its prob density function is:
$f_X(x)=\lambda e^{-\lambda x}$ for $x>0$, and 0 otherwise.
E(X) $=\frac{1}{\lambda}=\sigma$, V(X) $=\frac{1}{\lambda^2}$, $P(X>x)=e^{-\lambda x}$
Memoryless: $P(X>s+t|X>s)=P(X>t)$.
- **Normal Dist**: X follows a normal distribution with parameter $\mu\in\mathbb{R}$ and $\sigma>0$ denoted by $X\sim N(\mu,\sigma^2)$ if its prob density function is:
$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
E(X) $=\mu$, V(X) $=\sigma^2$
Sym about $x=\mu$, max value (at $x=\mu$) is $\frac{1}{\sqrt{2\pi}\sigma}$
- **Standard Normal**: Normal distribution with $\mu=0$ and $\sigma=1$, denoted with $Z\sim N(0,1)$. PDF:
$\phi(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$.
If $Y\sim N(\mu,\sigma^2)$, then $Z=\frac{Y-\mu}{\sigma}\sim N(0,1)$.
CDF $\Phi(z)=P(Z\leq z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{y^2}{2}}$ dy
- **Quantile**: Can look up the table to retrieve z such that $P(Z\leq z)=q$ for a given q.
- **Binomial → Normal**: $\mu=np,\sigma^2=np(1-p)$. X approx $\sim N(np,np(1-p))$ as $n\to\infty$.
Remember to apply continuity correction!
Good when $np>5$ and $n(1-p)>5$

Samplings

- **Sampling Dist of Sample Mean** $\bar{X}$: Let pop mean be μ, pop s.d. be σ and sample size be n.
Assumption: infinite pop or finite pop with replacement (aka prob same)
Distribution of $\bar{X}$: $E(\bar{X})=\mu_{\bar{X}}=E(X)$,
$V(\bar{X})=\frac{V(X)}{n}=\sigma_{\bar{X}}^2=\frac{\sigma_X^2}{n}$
- **Law of Large Number (LLN)**:
$P(|\bar{X}-\mu|>\epsilon)\to 0$ as $n\to 0$
- **CLT**$(N\to\infty)$: $\bar{X}\sim N(\mu,\frac{\sigma^2}{n}), Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim$N(0,1)
- **Diff of 2 samples**: If $n_1,n_2\geq 30$, $(\bar{X}_1-\bar{X}_2)$ follows normal dist. $E(\bar{X}_1-\bar{X}_2)$=$E(\bar{X}_1)-E(\bar{X}_2)$, $V(\bar{X}_1-\bar{X}_2)=V(\bar{X}_1)+V(\bar{X}_2)$
- **Gamma Fn**: $\Gamma(n)=\int_0^\infty x^{n-1}e^{-x}$ dx. For +ve integer n, $\Gamma(n)=(n-1)!$
- **Chi-Sq Dist**: $Y\sim\chi^2(n)$ (n-deg of freedom)
$f_Y(y)=\frac{1}{2^{n/2}\Gamma(n/2)}y^{n/2-1}e^{-y/2}$ for $y>0$
$E(Y)=n$, $V(Y)=2n$, $\chi^2(n)$ approx $\sim N(n,2n)$ for large n. If $X\sim N(\mu,\sigma^2)$, then $(\frac{X-\mu}{\sigma})^2\sim\chi^2(1)$
If $Y_1,Y_2,...,Y_k$ are indep χ^2 RV with $n_1,n_2,...,n_k$ deg of freedom, then $Y_1+..+Y_k\sim\chi^2(n_1+..+n_k)$
If $X_1,X_2,...,X_n$ be rand sam from a norm pop with mean μ and var σ^2, then
$Y=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$
$\chi^2(n;\alpha)$ is k satisfying $P(Y\geq k)=\alpha$, $Y\sim\chi^2(n)$

- **Samp Var**: Let $X_1,X_2,...X_n$ be rand sam from a pop. Sample variance $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$.
If S^2 is the variance of a rand sam of size n from a norm pop with var σ^2, then $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$
- **T-dist**: Let $Z\sim N(0,1)$, $U\sim\chi^2(n)$. If Z and U are indep, $T=\frac{Z}{\sqrt{U/n}}\sim t(n)$ (follows T-dist with n deg of freedom)
$f_T(t)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(n/2)}(1+\frac{t^2}{n})^{\frac{n+1}{2}}$ for $-\infty<t<\infty$
Sym about $x=0$; $\lim_{n\to\infty}f_T(t)=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$
$E(T)=0$ and $V(T)=n/(n-2)$ for $n>2$
$t_{n;\alpha}$ is k satisfying $P(T\geq k)=\alpha$, $T\sim t(n)$ Rand Sam from a norm pop: T=$\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
- **F-dist**: Let $U\sim\chi^2(n_1)$, $V\sim\chi^2(n_2)$ and U, V are indep. Then $F=\frac{U/n_1}{V/n_2}\sim F(n_1,n_2)$ (follows F-dist with (n_1, n_2) deg of freedom)
$f_F(x)=\frac{n_1^{n_1/2}n_2^{n_2/2}\Gamma(\frac{n_1+n_2}{2})}{\Gamma(n1/2)\Gamma(n2/2)}\frac{x^{(n_1/2+1)}}{(n_1x+n_2)^{(n_1+n_2)/2}}$
for $x>0$
With n_1 and n_2 samples,
$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$, σ are var of pop
If $F\sim F(n,m)$, then $1/F\sim F(m,n)$.
$F(n_1,n_2;\alpha)$ is k satisfying $P(F>k)=\alpha$,
$F\sim F(n_1,n_2)$. $F(n_1,n_2;1-\alpha)=1/F(n_2,n_1;\alpha)$

Estimation

- $f_X(x;\theta)$ is the pdf of RV X with an unknown parameter θ. Value of θ can be estimated.
- **Statistic**: A function of rand sample which does not depend on any unknown parameters. Eg: $max(X_1,X_2,...X_n)$
- **Point est**: Use a statistic to est the unknown param θ. The **statistic** will be called **point estimator**.
- **Interval est**: Two statistics $(\hat{\Theta}_L,\hat{\Theta}_R)$ consistutes an interval for which the prob of containing the unknown parameter θ can be determined.
- **Unbiased Estimator**: $E(\hat{\Theta})=\theta$ Eg: Unbiased: $E(\bar{X})$=μ, $E(S^2)$=σ^2 Biased: $E(T)$=$\frac{n-1}{n}\sigma^2\neq\sigma^2$
- **CI for** μ**, known** σ: $\bar{X}\pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ (See CLT)
If we want $P(|\bar{X}-\mu|\leq e)\geq 1-\alpha$, where e is margin of error, then $e\geq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, $n\geq(z_{\alpha/2}\frac{\sigma}{e})^2$
- **CI for** μ**, unknown** σ: $\bar{X}\pm t_{n-1;\alpha/2}(\frac{S}{\sqrt{n}})$
If $n>30$, $\bar{X}\pm z_{\alpha/2}(\frac{S}{\sqrt{n}})$
- **CI for** $\mu_1-\mu_2$**,** $\sigma_1^2\neq\sigma_2^2$:
$(\bar{X}_1-\bar{X}_2)\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ (known σ_1^2, σ_2^2)
n_1, $n_2\geq 30$: If σ_1^2, σ_2^2 dk, can subst with S_1^2, S_2^2.
- **CI for** $\mu_1-\mu_2$**,** $\sigma_1^2=\sigma_2^2$: If dk σ^2, use pooled sample variance $S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$
($\bar{X}_1$-$\bar{X}_2$) $\pm t_{n_1+n_2-2;\alpha/2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ (smal, nor)
($\bar{X}_1$-$\bar{X}_2$) $\pm z_{\alpha/2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ (n≥30)

- **CI for** $\mu_D=\mu_1-\mu_2$: $d_i=x_i-y_i$
$S_D^2=\frac{1}{n-1}\sum_{i=1}^{n}(d_i-\bar{d})^2$
$\bar{d}\pm t_{n-1;\alpha/2}(\frac{S_D}{\sqrt{n}})$ (smal, norm) $\bar{d}\pm z_{\alpha/2}(\frac{S_D}{\sqrt{n}})$ (big)
- **CI for** σ^2**, normal pop**:
$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ is a pt est of σ^2.
- **CI for** σ^2**, normal pop, known** μ:
$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{n;\alpha/2}}<\sigma^2<\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{n;1-\alpha/2}}$
- **CI for** σ^2**, normal pop, unknown** μ:
$\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}}<\sigma^2<\frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}$, s = sample var
- **CI for** σ: To get σ, just take sq root of range.
- **CI for** $\frac{\sigma_1^2}{\sigma_2^2}$**: norm pop, unknown** $\mu_1\mu_2$:
$\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1;n_2-1;\alpha/2}}<\frac{\sigma_1^2}{\sigma_2^2}<\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1;n_2-1;1-\alpha/2}}=$
$\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1;n_2-1;\alpha/2}}<\frac{\sigma_1^2}{\sigma_2^2}<\frac{S_1^2}{S_2^2}F_{n_2-1;n_1-1;\alpha/2}$
Similarly to find $\frac{\sigma_1}{\sigma_2}$ just take sqrt.

Testing

- **Type1 err**: Rej H_0 when H_0 is true. P(type1) = α = Significance Level
- **Type2 err**: Didn't rej H_0 when H_0 is false. P(type2) = P(H_0 accepted — H_1) = β.
- **Power**=$1-\beta$
- **Rej/Crit Region**: Area that causes H_0 to be rej.
- **p-value**: Prob of obtaining a more extr result than the sam mean/var. Rmb to multiply by 2 for 2TT.
- **Two-tail test**: $H_0:\mu=\mu_0$ vs $H_1:\mu\neq\mu_0$
Can calculate CI and reject if not in CI.
- **Test for** μ**, known** σ: $Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$
Reject if $|Z|>z_{\alpha/2}$ (2TT) or $|Z|>z_\alpha$ (1TT).
- **Test for** μ**, dk** σ: $T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim T(n-1)$ Reject if $|T|>t_{n-1;\alpha/2}$ (2TT) or $|T|>t_{n-1;\alpha}$ (1TT).
- **Test for** $\mu_1-\mu_2$: $Z=\frac{(\bar{X}_1-\bar{X}_2)-\mu_0}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$
$Z=\frac{(\bar{X}_1-\bar{X}_2)-\mu_0}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\sim N(0,1)$ (big n, dk σ)
- **Test for** $\mu_1-\mu_2$**,** $\sigma_1^2=\sigma_2^2$:
$T=\frac{(\bar{X}_1-\bar{X}_2)-\mu_0}{S_p\sqrt{1/n_1+1/n_2}}\sim T(n_1+n_2-2)$
- **Test for** μ_D, $D_i=X_i-Y_i$: $T=\frac{\bar{D}-\mu_{D,0}}{S_D/\sqrt{n}}$
$T\sim t(n-1)$ (n<30, norm); $T\sim N(0,1)$ (n≥30)
- **Test for** σ^2: $\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$
$\sigma^2>\sigma_0^2$: rej if $\chi^2>\chi^2_{n-1;\alpha}$
$\sigma^2<\sigma_0^2$: rej if $\chi^2<\chi^2_{n-1;1-\alpha}$
$\sigma^2\neq\sigma_0^2$: rej if $\chi^2<\chi^2_{n-1;1-\alpha/2}$ or $\chi^2>\chi^2_{n-1;\alpha/2}$
- **Test for** $H_0:\sigma_1^2=\sigma_2^2$:
$F=\frac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$
$\sigma_1^2>\sigma_2^2$: rej if $F>F_{n_1-1;n_2-1;\alpha}$
$\sigma_1^2<\sigma_2^2$: rej if $F<F_{n_1-1;n_2-1;1-\alpha}$
$\sigma_1^2\neq\sigma_2^2$: rej if $F<F_{n_1-1;n_2-1;1-\alpha/2}$ or $F>F_{n_1-1;n_2-1;\alpha/2}$