ST2334 Cheatsheet AY23/24 —— @Jin Hang

Observation: Recording of information, numerical or categorical
Statistical Exp: Procedure that generates a set of observations
Sample Space: Set of all possible outcomes of a statistical experiment, represented by the symbol S.
Sample Points: Every outcome in a sample space.
Events: Subset of sample space.
Simple/Compound Event: Exactly one/More than one outcome or sample point.
Sure/null Event: Sample space/Event with no outcomes.

Operations on Events

- **Complement**: A' Elements not in A
- **Mutually Exclusive:(互斥)** $A \cap B = \emptyset$
- **Independent(独立):** $A \perp B \Rightarrow P(AB) = P(A) \cdot P(B)$
  1. **独立和互斥的关系:** 独立不互斥, 互斥不独立
  2. S and $\emptyset$ are independent of any other event.
  3. If $A \perp B$, then $A \perp B', A' \perp B$, and $A' \perp B'$.
  4. $A \perp B \Rightarrow P(A) = P(A|B)$ & $P(B) = P(B|A)$
  5. $P(A) = 0/1 \Rightarrow A$与任意事件B独立
  6. $0 < P(A), P(B) < 1$, 若AB互斥/包含 $\Rightarrow$ 不独立
- **不相关:** $cov(X,Y) = 0$, 独立必不相关, 不相关未必独立
- **Union:** $A \cup B$ contains A or B or both elements
- **Intersection:** $A \cap B$ contains elements common to both

1. $A \cap A' = \emptyset$
2. $A \cap \emptyset = \emptyset$
3. $A \cup A' = S$
4. $(A')' = A$
5. $(A \cap B)' = A' \cup B'$
6. $(A \cup B)' = A' \cap B'$
7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
8. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
9. $A \cup B = A \cup (B \cap A')$
10. $A = (A \cap B) \cup (A \cap B')$

$A \subset B$ if all elements in event A are in event B, if $A \subset B$ and $B \subset A$ then $A = B$. We assume **contained** means proper subset.

Permutations and Combinations

- P: Arrange r objects from n objects where $r \le n$, ${}_nP_r = \frac{n!}{(n-r)!}$
- P: Number of ways around a circle $= (n-1)!$
- C: Select r from n without regard order: $\binom{n}{r} = {}_nC_r = \frac{n!}{r!(n-r)!}$
- $\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$
- $\binom{n}{r} = \binom{n}{n-r}$
- $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n$

Permutations with Sets of Indistinguishable Objects

Suppose collection consists n obj. with its n_k are of type k and are distinct from each other, $n_1 + n_2 + \ldots + n_k = n$. Then the number of distinguishable permutations of the n objects is

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-n_2-\ldots-n_{k-1}}{n_k} = \frac{n!}{n_1!n_2!n_3!\cdots n_k!}$$

常用方法	Order Matters	Order Does Not Matter
Repetition	n^k	$\binom{k+n-1}{k}$
No Repetition	$P(n,k)$	$\binom{n}{k}$

方程模型 How many solutions are there to the give equations?
(a) $x_1 + x_2 + x_3 = 20$, each x_i is a nonnegative integer.
Solution. $n = 3, r = 20$.

$$\binom{r+(n-1)}{r} = \binom{20+2}{20} = \binom{22}{20} = \frac{22 \cdot 21}{2} = 231$$

(b) $x_1 + x_2 + x_3 = 20$, each x_i is a positive integer.
Solution. $\Leftrightarrow y_1 + y_2 + y_3 = 17$, each y_i is a nonnegative integer

$$\binom{r+(n-1)}{r} = \binom{17+2}{17} = \binom{19}{17} = \frac{19 \cdot 18}{2} = 171$$

Probability

- **Conditional:** $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$, if $\Pr(A) \neq 0$
- **Multiplicative:** $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B|A) = \mathrm{P}(B)\mathrm{P}(A|B)$
- **LoTP:** $\mathrm{P}(B) = \sum_{i=1}^n \mathrm{P}(B \cap A_i) = \sum_{i=1}^n \mathrm{P}(A_i)\mathrm{P}(B|A_i)$
- **Bayes:** $\mathrm{P}(A_k|B) = \frac{\mathrm{P}(A_k)\mathrm{P}(B|A_k)}{\sum_{i=1}^n \mathrm{P}(A_i)\mathrm{P}(B|A_i)}$

Probability Axioms

- $P(A \cup B) = P(A) + P(B) \le 1$
- $P(AB) \le P(A)$ or $P(B)$
- $P(A) = P(A \cap B) + P(A \cap B')$
- $A \subset B \Rightarrow P(A) \le P(B)$
- If $P(A) \neq 0, A \perp B$ if and only if $P(B \mid A) = P(B)$

Properties of Independent Events

1. $\mathrm{P}(B|A) = \mathrm{P}(B)$ and $\mathrm{P}(A|B) = \mathrm{P}(A)$
2. A and B cannot be mutually exclusive if they are independent, supposing $\mathrm{P}(A), \mathrm{P}(B) > 0$
3. A and B cannot be independent if they are mutually exclusive
4. Sample space S and empty set $\emptyset$ are independent of any event
5. If $A \subset B$, then A and B are dependent unless $B = S$.

Random Var. Range Space: The range space of X is the set of real numbers $R_X = \{x | x = X(s), s \in S\}$. The set $X = x = s \in S : X(s) = x$ is a subset of S.

Probability Mass Function

Discrete Random Variable: $0 \le f(x) = \begin{cases} P(X = x), x \in R_X \\ 0, x \notin R_X \end{cases}$

Continuous Random Variable: The **probability density function** (p.d.f.) $f(x)$ of a continuous random variable must satisfy the following conditions

- $f(x) \ge 0$ for all $x \in R_X$ and $f(x) = 0$ for $x \notin R_X$ i.e. $\mathrm{P}(A) = 0$ does not imply $A = \emptyset$
- $\int_{R_X} f(x)dx = 1 \Rightarrow \int_{-\infty}^{+\infty} f(x)dx = 1$
- For any $(c,d) \subset R_X, c < d$, $\mathrm{P}(c \le X \le d) = \int_c^d f(x)dx$
- Specially $\mathrm{P}(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$
- $P(a < X < b) = P(a < X \le b) = P(a \le X < b) = P(a \le X \le b) = \int_a^b f(x)dx$

Cumulative Distribution Function c.d.f $F(x) = \mathrm{P}(X \le x)$

- $F(x) = \int_{-\infty}^x f(t)dt$
- $f(x) = \frac{dF(x)}{dx}$(连续点处)

Properties:
1. $F(x)$ is **non-decreasing**. 2. **one-to-one** correspondence
3. $F(x)$右连续, $F(a+0) = \lim_{x \to a^+} F(x) = F(a)$

Range: The ranges of $F(x)$ and $f(x)$ satisfy
- $0 \le F(x) \le 1$;
- for discrete distributions, $0 \le f(x) \le 1$;
- for cont. distributions, $f(x) \ge 0$, but **not necessary** $f(x) \le 1$.

利用CDF求概率问题

- $P\{a \le X \le b\} = \mathrm{P}(X \le b) - \mathrm{P}(X < a) = F(b) - F(a-0)$
- $P\{a < X \le b\} = F(b) - F(a)$
- $P\{a < X < b\} = F(b-0) - F(a)$
- $P\{a \le X < b\} = F(b-0) - F(a-0)$

Expectation / Mean

- $\mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$; $E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
- $E(aX+b) = aE(X) + b$
- $E(X+Y) = E(X) + E(Y)$

Variance $g(x) = (x - \mu_X)^2$, leads us to the definition of variance.
$\sigma_X^2 = V(X) = E[g(x)] = E[(X - \mu_X)^2] =$

$$\begin{cases} \sum_x (x-\mu_X)^2 f_X(x), & \text{if } X \text{ is discrete,} \\ \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)dx, & \text{if } X \text{ is continuous.} \end{cases}$$

Properties
- $\forall X, V(X) \ge 0$. "=" iff. $P(X = E(X)) = 1$ (X is constant).
- $\sigma_X = \sqrt{V(X)}$ (**Standard deviation**)
- $V(X) = E(X^2) - [E(X)]^2$
- $V(aX+b) = a^2V(X)$
- $V(aX+bY) = a^2V(X) + b^2V(Y) + 2abCov(X,Y)$, if X,Y independent, $Cov(X,Y) = 0$

2-D random variable (X,Y) is a two-dimensional random var., where X, Y are functions assigning a real number to each $s \in S$.
Range Space: $R_{X,Y} = \{(x,y) | x = X(s), y = Y(s), s \in S\}$

Joint Probability Function

Discrete: $f_{X,Y}(x_i, y_j) = \mathrm{P}(X = x_i, Y = y_j)$

1. $f_{X,Y}(x_i, y_j) \ge 0$ for all $(x_i, y_j) \in R_{X,Y}$.
2. $\sum_{i=1}^\infty \sum_{j=1}^\infty f_{X,Y}(x_i, y_j) = 1$.

Continuous
1. $P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(x,y)dydx$
2. $f_{X,Y}(x,y) \le 0$ for all $(x,y) \in R_{X,Y}$.
3. $\iint_{(x,y)\in R_{X,Y}} f_{X,Y}(x,y)dxdy = \int_\infty^\infty \int_\infty^\infty f_{X,Y}(x,y)dxdy = 1$.

Marginal Probabilty Densities

- **Discrete:** $f_X(x) = \sum_y f_{X,Y}(x,y)$ and $f_Y(y) = \sum_x f_{X,Y}(x,y)$
- **Cont:** $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$, $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$

Conditional Probabilty Densities

The conditional distribution of Y given that $X = x$ is given by $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$, iff. $f_X(x) > 0$, for each x within the range of X. Flip the variables for X given $Y = y$.

1. Consider y as the variable (and x as a fixed value), $f_{Y|X}(y|x)$ is a **probability function**.
2. $f_{Y|X}(y|x)$ is NOT a probability function for $x \Rightarrow$ No need for $\int_{-\infty}^{+\infty} f_{Y|X}(y|x)dx = 1$ or $\sum_x f_{Y|X}(y|x) = 1$
3. if $f(x) > 0 \Rightarrow f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x)$
4. For Cont. $P(Y \le y | X = x) = \int_{-\infty}^y f_{Y|X}(y|x)dy$
   $E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy$

Independent Random Variables

Random variables X and Y are independent if and only if $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all x, y, extendable to n variables.
$\Rightarrow$ Implying $R_{X,Y} = \{(x,y) | x \in R_X; y \in R_Y\}$
$\Rightarrow$ If $f_X(x) > 0$ then $f_{Y|X}(y|x) = f_Y(y)$
To Check independence, we have $f_{X,Y}(x,y) = C \times g_1(x) \times g_2(y)$

Expectation

1. Discrete, $E(g(X,Y)) = \sum_x \sum_y g(x,y) f_{X,Y}(x,y)$
2. Continous $E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dydx$
3. $E(XY) = \int\int xy(f(x,y))dydx$
4. $X \perp Y \Rightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y) \Rightarrow E(XY) = E(X)E(Y)$

Covariance

1. $cov(X,Y) = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y)$
2. $cov(X,Y) = E[(X - E(X)(Y - E(Y)] = E(XY) - \mu_X\mu_Y$.
3. $X \perp Y \Rightarrow cov(X,Y) = 0$. and $cov(X,Y) = 0 \nRightarrow X \perp Y$
4. $cov(X,Y) = cov(Y,X)$
5. $cov(X+b, Y) = cov(X,Y)$
6. $cov(aX, Y) = acov(X,Y)$
7. $cov(aX+b, cY+d) = ac \cdot cov(X,Y)$.
8. $V(X+Y) = V(X) + V(Y) + 2cov(X,Y)$
9. $V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab \cdot cov(X,Y)$.
10. If $X \perp Y$, $V(aX - bY) = a^2V(X) + b^2V(Y)$ $(cov = 0)$
11. $V(X - Y) = V(X) + V(Y)$

Distributions

Discrete Uniform Distribution: $f_X(x) = \frac{1}{k}$
- $E(X) = \frac{1}{k}\sum_{i=1}^k x_i$
- $V(X) = \frac{1}{k}\sum_{i=1}^k x_i^2 - \mu_X^2$

Bernoulli random variable: $f_X(x) = p^x(1-p)^{1-x}$
- $E(X) = p$
- $V(X) = p(1-p) = pq$

Binomial random variable: $X \sim \mathrm{Bin}(n,p)$
- $P(X = x) = \binom{n}{x}p^x(1-p)^{n-x}$
- $E(X) = np$
- $V(X) = np(1-p) = npq$

Negative Binomial distribution: $X \sim \mathrm{NB}(k,p)$,
- $f_X(x) = P(X = x) = \binom{x-1}{k-1}p^k(1-p)^{x-k}$
- $E(X) = \frac{k}{p}$
- $V(X) = \frac{(1-p)k}{p^2}$

Geometric distribution: Numbers need until first success occur. $f_X(x) = P(X = x) = (1-p)^{x-1}p$, $E(X) = \frac{1}{p}$, $V(X) = \frac{1-p}{p^2}$
Poisson random variable: k is num(occurrences) of events

- $f_X(k) = P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$
- $E(X) = \lambda$ • $V(X) = \lambda$ • $X \sim \text{Poisson}(\lambda)$
- $n \to \infty, p \to 0$, $\text{Bin}(n, p) \to \text{Poisson}(np)$ if $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$.

Uniform distribution: $X \sim U(a,b)$
- $f_X(x) = \frac{1}{b-a}, a \leq x \leq b$ • $F_X(x) = \frac{x-a}{b-a}, a \leq x \leq b$
- $E(X) = \frac{a+b}{2}$ • $V(X) = \frac{(b-a)^2}{12}$

Exponential distribution: $X \sim \text{Exp}(\lambda)$
- $f_X(x) = \lambda e^{-\lambda x}$, if $x \geq 0$ • $F_X(x) = 1 - e^{-\lambda x}$
- $E(X) = \frac{1}{\lambda}$ • $V(X) = \frac{1}{\lambda^2}$
- **Memoryless:** $P(X > s+t \mid X > s) = P(X > t)$
- $X_1, X_2, \ldots, X_n$ independent and identically distributed $\text{Exp}(\lambda)$ distributions. $X = \min\{X_1, X_2, \ldots, X_n\} \Rightarrow X \sim \text{Exp}(n\lambda)$
  $P(X > x) = P(\min\{X_1, X_2, \ldots, X_n\} > x) =$
  $P(X_1 > x, X_2 > x, \ldots, X_n > x) =$
  $P(X_1 > x) P(X_2 > x) \cdots P(X_n > x) = e^{-\lambda x} \cdots e^{-\lambda x} = e^{-n\lambda x}$

Normal distribution: $X \sim N(\mu, \sigma^2)$
- $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$
- $\phi(z) = f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$
- Two normal curves are identical in shape if they have same σ^2
- As σ increases, the curve flattens; and vice versa.
- $x_1 < X < x_2 \Longleftrightarrow \frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{x_2-\mu}{\sigma}$
- **C.D.F.** $\Phi(z) = \int_{-\infty}^{z} \phi(t)\text{d}t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\text{d}t$
- $X \sim N(\mu, \sigma^2)$ $P(x_1 < X < x_2) = \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right)$
- $P(Z \geq 0) = P(Z \leq 0) = \Phi(0) = 0.5$
- $\Phi(z) = P(Z \leq z) = P(Z \geq -z) = 1 - \Phi(-z)$
- $P\{|X| \leqslant a\} = 2\Phi(a) - 1, a > 0$

Property
- $E(X) = \mu$ • $V(X) = \sigma^2$
- $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$
- $Z \sim N(0,1) \Rightarrow -Z \sim N(0,1) \Rightarrow \sigma Z + \mu \sim N(\mu, \sigma^2)$
- $\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$
- **αth (upper) quantile** $P(X \geq x_\alpha) = \alpha$
- **Approximation:** For $np > 5$ and $n(1-p) > 5$，$X \sim \text{Bin}(n,p)$, as $n \to \infty$, $Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1)$.

Continuity Correction
- $P(X = k) \approx P(k - \frac{1}{2} < X < k + \frac{1}{2})$
- $P(a \leq X \leq b) \approx P(a - \frac{1}{2} < X < b + \frac{1}{2})$
- $P(a < X \leq b) \approx P(a + \frac{1}{2} < X < b + \frac{1}{2})$
- $P(a \leq X < b) \approx P(a - \frac{1}{2} < X < b - \frac{1}{2})$
- $P(a < X < b) \approx P(a + \frac{1}{2} < X < b - \frac{1}{2})$

Sample Statistics: Sample resembles Population!
Population parameters: Values Computed (μ, σ, p)
Statistic: Suppose a random sample of n, $(X_1, \ldots, X_n)$ has been taken. A function of $(X_1, \ldots, X_n)$ is called a **statistic**.
1. **Sample mean:** $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ ($\bar{X}, S$ are **Random Var.**)
2. **Sample variance:** $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$

Mean and Variance of $\bar{X}$:
- $\mu_{\bar{X}} = E(\bar{X}) = \mu_X$ • $\sigma_{\bar{X}}^2 = V(\bar{X}) = \frac{\sigma_X^2}{n}$
- $E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] = \sigma^2$
- $\bar{X}$ estimates μ_X, $n \uparrow$. $\sigma_X^2/n \downarrow \Rightarrow \mu_X \to \bar{X}$.

大数定律(LLN): $P(|\bar{X} - \mu| > \varepsilon) \to 0$ as $n \to \infty$

中心极限定理(CLT):
$n \to \infty \Rightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \to Z \sim N(0,1) \Leftrightarrow \bar{X} \to N\left(\mu, \frac{\sigma^2}{n}\right)$
1. $\lim_{n\to\infty} P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$
2. $X_1, X_2, \ldots, X_n$ independent and $N(\mu, \sigma^2)$, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ or $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ regardless n.

χ^2 Distribution: Let $Z_1, \ldots, Z_n$ be n independent and **standard normal** random variables, then $Z_1^2 + \cdots + Z_n^2$ is called a χ^2 random variable with n degrees of freedom.

Properties
1. $Y \sim \chi^2(n) \Rightarrow E(Y) = n$ and $V(Y) = 2n$.
2. $X \sim N(0,1) \Rightarrow X^2 \sim \chi^2(1) \Rightarrow E(X^2) = 1, V(X^2) = 2$
3. For large n, $\chi^2(n)$ is approximately $N(n, 2n)$.
4. $Y_1 \sim \chi^2(m)$ and $Y_2 \sim \chi^2(n) \Rightarrow Y_1 + Y_2 \sim \chi^2(m+n)$
5. χ^2 distribution curve is determined by n with long right tail.
6. If S^2 is the variance of a random sample of size n taken from a normal population having the variance σ^2, then $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2}$ has a $\chi^2(n-1)$ distribution.

t-distribution: $Z \sim N(0,1)$, $U \sim \chi^2(n)$. If Z and U independent, then $T = \frac{Z}{\sqrt{U/n}} \sim t(n)$

Properties
1. The t-distribution approaches $N(0,1)$ as $n \to \infty$ i.e. $n \geq 30$
2. If $T \sim t(n)$, then $E(T) = 0$ and $V(T) = n/(n-2)$ for $n > 2$.
3. t-分布概率密度$f(x)$为偶函数, $n \to \infty$时接近$N(0,1)$分布
4. $X_i \sim N(\mu, \sigma^2)$, $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

F-distribution: Suppose $U \sim \chi^2(m)$ and $V \sim \chi^2(n)$ are independent. Then $F = \frac{U/m}{V/n} \sim F(m,n)$

Properties
1. If $X \sim F(m,n)$, then $E(X) = \frac{n}{n-2}$ and $V(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$
2. If $F \sim F(n,m)$, then $1/F \sim F(m,n)$.
3. For $P(F > F(m,n;\alpha)) = \alpha$, $F(m,n;1-\alpha) = 1/F(n,m;\alpha)$
4. If $T \sim t(n)$, then $T^2 \sim F(1,n)$.

If $X_1, \ldots, X_n$ independent & identically distributed $X_i \sim N(0,1)$, $\bar{X}$ is the sample mean and S^2 is the sample variance, then
- $n\bar{X} \sim N(0,n)$.
- $(n-1)S^2 \sim \chi_{n-1}^2$.
- $\frac{\sqrt{n}\bar{X}}{S} \sim t_{n-1}$.
- $\frac{(n-1)X_1^2}{\sum_{i=2}^{n} X_i^2} \sim F(1, n-1)$.

Estimation
Unbiased Estimator Let $\hat{\Theta}$ be an **unbiased** estimator of θ. Then $\hat{\Theta}$ is a random variable based on the sample s.t. $E(\hat{\Theta}) = \theta$
- $\bar{X}$ is a good estimator of μ
- $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu_X = \mu_X$
- $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$, $E(S^2) = \sigma^2$
- $\text{var}(\bar{X}) = \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \text{var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_X^2 = \frac{\sigma_X^2}{n}$

z_α: Number with upper-tail prob. of α s.t. $P(Z > z_\alpha) = \alpha$.
Maximum Error of Estimate: $\bar{X} \neq \mu \Rightarrow \bar{X} - \mu$ measures difference between estimator and the true value of the parameter. If population is normal or n is large, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ follows a standard normal or an approximately standard normal distribution.
$P\left(\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = P\left(|\bar{X} - \mu| \leq z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$
$\Rightarrow E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Determine Sample Size: $z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq E_0 \Rightarrow n \geq \left(\frac{z_{\alpha/2} \cdot \sigma}{E_0}\right)^2$

Interval Estimator: For (a,b) you are fairly certain the parameter of interest lies in, quantified by confidence level $(1-\alpha)$ s.t. $P(a < \mu < b) = 1 - \alpha$. (a,b) is $(1-\alpha)$ confidence interval.
- $(1-\alpha)$ confidence interval can be written as $\bar{X} \pm E$.
- $\bar{X} \pm E$ has probability $(1-\alpha)$ of containing μ
- Once computed, μ is either in it or not $\Rightarrow$no more randomness.
- n is large when $n \geq 30$

Comparing Two Population: Confidence Intervals for $\mu_1 - \mu_2$
- **Independent samples**: complete randomization.
- **Matched pairs samples**: randomization between pairs.

Pooled estimator(S_p^2): σ^2 can be estimated by the **pooled sample variance** $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$ with S_1^2 and S_2^2 being the sample variances of the first and second samples respectively.
Roughly assume equal variance if $1/2 \leq S_1/S_2 \leq 2$ not sensitive to small difference between population var.
Paired Data: For $(X_1, Y_1), \ldots, (X_n, Y_n)$
- X_i and Y_i are dependent. • (X_i, Y_i) are independent
- Define random sample $D_i = X_i - Y_i, \mu_D = \mu_1 - \mu_2$.
- **Small** and Normal: $\bar{d} \pm t_{n-1;\alpha/2} \cdot \frac{s_D}{\sqrt{n}}$; **Large:** $\bar{d} \pm z_{\alpha/2} \cdot \frac{s_D}{\sqrt{n}}$

Hypothesis Tests
- **Null**: Try to it's false (Type I may happen reject null.), makes an assertion that a parameter equals to some constant.
- **Ater.**: Prove to be true, against **null.** Reject $H_0 \Rightarrow$ Concl. H_1 Type II occur if do not reject null
- Reject null. $\Rightarrow$ enough evidence to support alternative.

Type I/II Error	Do not reject H_0	Reject H_0
H_0 is true	Correct Decision	Type I error
H_0 is false	Type II error	Correct Decision

- The Type I error: serious $\to$ control P(Type I)
- Thus prior to conducting a hypothesis test, we set the significance level α to be small, typically at $\alpha = 0.05$ or 0.01
- Did not "prove" that H_0 is true $\Rightarrow$ Not accept

p-value(observed level of significance): Probability of obtaining a test statistic at least as extreme ($\leq$ or $\geq$) than the observed sample value, given H_0 is true.

$\mu \neq \mu_0$	$\mu < \mu_0$	$\mu > \mu_0$
$P(\lvert Z\rvert > \lvert z\rvert)$	$P(Z < -\lvert z\rvert)$	$P(Z > \lvert z\rvert)$

- p-value $< \alpha$, reject H_0; p-value $\geq \alpha$, do not reject H_0

H_1	Rejection Region	p-value
$\mu_1 - \mu_2 > \delta_0$	$z > z_\alpha$	$P(Z > \lvert z\rvert)$
$\mu_1 - \mu_2 < \delta_0$	$z < -z_\alpha$	$P(Z < -\lvert z\rvert)$
$\mu_1 - \mu_2 \neq \delta_0$	$z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$	$2P(Z > \lvert z\rvert)$

Level of significance:
$\alpha = P($ Type I error $) = P($ Reject $H_0 \mid H_0$ is true $)$
Power of Test: $1 - \beta = P($ Reject $H_0 \mid H_0$ is false $)$, where $\beta = P($ Type II error $) = P($ Do not reject $H_0 \mid H_0$ is false $)$
$\alpha \uparrow \beta \downarrow \quad \alpha + \beta \neq 1$

Rejection Region:
- $H_1 : \mu \neq \mu_0$, $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
- $H_1 : \mu < \mu_0$, $z < -z_\alpha$
- $H_1 : \mu > \mu_0$, $z > z_\alpha$

Step 1: Set the null and alternative.
Step 2: Set $\alpha = 0.05$
Step 3: Use test with test statistics, determine rejection region
Step 4: Calculate observed value use Step 3 distribution
Step 5: Reject/Not Reject

Appendix 标准积分表

- $\int \frac{1}{ax+b}dx = \frac{1}{a}\ln|ax+b| + C$
- $\int \tan(ax+b)dx = \frac{1}{a}\ln|\sec(ax+b)| + C$
- $\int \sec(ax+b)dx = \frac{1}{a}\ln|\sec(ax+b)+\tan(ax+b)| + C$
- $\int \csc(ax+b)dx = -\frac{1}{a}\ln|\csc(ax+b)+\cot(ax+b)| + c$
- $\int \cot(ax+b)dx = -\frac{1}{a}\ln|\csc(ax+b)| + C$
- $\int \sec^2(ax+b)dx = \frac{1}{a}\tan(ax+b) + C$
- $\int \csc^2(ax+b)dx = -\frac{1}{a}\cot(ax+b) + C$
- $\int \sec(ax+b)\tan(ax+b)dx = \frac{1}{a}\sec(ax+b) + C$
- $\int \csc(ax+b)\cot(ax+b)dx = -\frac{1}{a}\csc(ax+b) + C$
- $\int \frac{1}{a^2+(x+b)^2}dx = \frac{1}{a}\tan^{-1}\left(\frac{x+b}{a}\right) + C$
- $\int \frac{1}{\sqrt{a^2-(x+b)^2}}dx = \sin^{-1}\left(\frac{x+b}{a}\right) + C$
- $\int \frac{-1}{\sqrt{a^2-(x+b)^2}}dx = \cos^{-1}\left(\frac{x+b}{a}\right) + C$
- $\int \frac{1}{a^2-(x+b)^2}dx = \frac{1}{2a}\ln\left|\frac{x+b+a}{x+b-a}\right| + C$
- $\int \frac{1}{(x+b)^2-a^2}dx = \frac{1}{2a}\ln\left|\frac{x+b-a}{x+b+a}\right| + C$
- $\int \frac{1}{\sqrt{(x+b)^2+a^2}}dx = \ln\left|(x+b)+\sqrt{(x+b)^2+a^2}\right| + C$
- $\int \frac{1}{\sqrt{(x+b)^2-a^2}}dx = \ln\left|(x+b)+\sqrt{(x+b)^2-a^2}\right| + C$
- $\int \sqrt{a^2-x^2}dx = \frac{x}{2}\sqrt{a^2-x^2} + \frac{a^2}{2}\sin^{-1}\frac{x}{a} + C$
- $\int \sqrt{x^2-a^2}dx = \frac{x}{2}\sqrt{x^2-a^2} - \frac{a^2}{2}\ln\left|x+\sqrt{x^2-a^2}\right| + C$

三角恒等变换

半角公式

- $\cos^2\alpha = \frac{1+\cos 2\alpha}{2}$
- $\sin^2\alpha = \frac{1-\cos 2\alpha}{2}$
- $\tan 2\alpha = \frac{2\tan\alpha}{1-\tan^2\alpha}$
- $\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}$
- $\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}$
- $\tan\frac{\alpha}{2} = \frac{\sin\alpha}{1+\cos\alpha} = \frac{1-\cos\alpha}{\sin\alpha}$

和差化积公式

- $\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$
- $\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}$
- $\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$
- $\cos\alpha - \cos\beta = -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}$

积化和差公式

- $\sin\alpha\cos\beta = \frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)]$
- $\cos\alpha\sin\beta = \frac{1}{2}[\sin(\alpha+\beta)-\sin(\alpha-\beta)]$
- $\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)]$
- $\sin\alpha\sin\beta = -\frac{1}{2}[\cos(\alpha+\beta)-\cos(\alpha-\beta)]$

万能公式

- $\sin\alpha = \frac{2\tan\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}$
- $\cos\alpha = \frac{1-\tan^2\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}$
- $\tan\alpha = \frac{2\tan\frac{\alpha}{2}}{1-\tan^2\frac{\alpha}{2}}$
- $\tan(\alpha+\beta) = \frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}$