ST2334 Cheatsheet 19/20 S1 Finals

Introduction

- Sample Space: S is the set of all possible outcomes
  - eg. For rolling 2 dice: $S = \{(1,1), ...(6,5), (6,6)\}$
- Sample Point: Any outcome in the sample space S
- Event: Any subset E of the sample space
- Sure Event: the sample space itself
- Null Event: empty set $\emptyset$

Counting

Choose k from n	Order Matters	Not Matter
With Replacement	n^k	$\binom{n+k-1}{k}$
Without Replacement	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

- In a circle: $(n-1)!$

Probability

Inclusion-Exclusion Principle

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - [P(A \cap B) + P(A \cap C) + P(B \cap C)] + P(A \cap B \cap C)$

Independent Events

- $P(A \cap B) = P(A) \times P(B)$ (use this to prove)
- $P(A|B) = P(A)$
- $P(A) = P(A \cap B) + P(A \cap B^c)$

Mutually Exclusive Events

- $P(A \cap B) = 0$ (B cannot happen if A happens)
- $P(A|B) = 0$
- $P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$

2 non-trivial ($P > 0$) events can only be independent, or mutually exclusive, or neither, but **never both**

Conditional Probability

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

De Morgan's Law

- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$

Partition

- If $B_1, B_2, ..., B_n$ are **mutually exclusive** and **exhaustive** (they are disjoint and their union $= S$), then $B_1, B_2, ..., B_n$ is a partition of S

Law of Total Probability (Bayes' Formula 1)

If $B_1, B_2, \ldots, B_n$ is a partition of S:

- $P(A) = \sum_{i=1}^n P(B_i \cap A) = \sum_{i=1}^n P(B_i)P(A|B_i)$

With extra conditioning:

- $P(A|C) = \sum_{i=1}^n P(A|B_i \cap C)P(B_i|C) = \sum_{i=1}^n P(A \cap B_i|C)$

Special case when B and B^c are the partitions:

- $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$
- $P(A) = P(A \cap B) + P(A \cap B^c)$

Bayes' Theorem

Let $B_1, \ldots, B_n$ be a partition of S. $\forall k \in 1, \ldots, n$,

- $P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^n P(B_i)P(A|B_i)}$

Discrete Random Variables

Probability Mass Function ($f_X(x)$)

- Probability that a discrete random variable $= x$
- Given by $f_X(x) = P(X = x)$
- When asked to find PMF: find $\forall x$, $P(X = x)$

Properties:

1. $0 \le f_X(x) \le 1$
2. $\sum_x f_X(x) = 1$
3. $P(X \in E) = \sum_{x \in E} f_X(x)$

Cumulative Distribution Function ($F_X(x)$)

- Probability that a discrete random variable is $\le x$
- $F_X(x) = P(X \le x) = \sum_{t \le x} P(X = t)$

Properties:

1. $F_X(x)$ is a non-decreasing function of x
2. $0 \le F_X(x) \le 1$

Continuous Random Variables

- $P(X = x) = 0$, so $P(a < X < b) = P(a \le X \le b)$
- $P(a < X < b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a)$

Probability Density Function ($f_X(x)$)

f_X is PDF of the continuous random variable X iff

1. $\forall x,\ f_X(x) \ge 0$
2. $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

Cumulative Distribution Function ($F_X(x)$)

- $F_X(x) = P(X \le x) = \int_{-\infty}^x f_X(t)\,dt$

Properties:

1. $F_X(x)$ is a non-decreasing function of x
2. $\lim_{x \to -\infty} F_X(x) = 0$ **AND** $\lim_{x \to \infty} F_X(x) = 1$

Mean & Variance

Mean ($E(X)\,|\,\mu_X$)

- **Discrete**: $E(X) = \sum_x xP(X = x)$
- **Continuous**: $E(X) = \int_{-\infty}^{\infty} xf_X(x)\,dx$
- $E[g(X)] = \sum_x g(x)f_X(x)$ **OR** $\int_{-\infty}^{\infty} g(x)f_X(x)\,dx$
  - eg. k^{th} moment: $E(X^k) = \sum_x (x)^k P(X = x)$

Properties:

1. $E(aX + bY + c) = aE(X) + bE(Y) + c$

Variance ($V(X)\,|\,\sigma_X^2$)

- **Discrete**: $V(X) = \sum_x (x - \mu_X)^2 f_X(x)$
- **Continuous**: $V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx$
- $\sigma_X = SD(X) = \sqrt{V(X)}$

Properties:

1. $V(X) \ge 0$
2. $V(X) = E(X^2) - [E(X)]^2$
3. $V(X) = 0 \implies P(X = \mu_X) = 1$ (data no spread)
4. $V(a + bX) = b^2V(X)$

Chebyshev's Inequality

If a random variable X has mean, μ, and SD, σ, the probability of getting a value which deviates from μ by at least $k\sigma$ is at most $\frac{1}{k^2}$

- $P(|X - \mu| > k\sigma) \le \frac{1}{k^2}$
- $P(|X - \mu| \le k\sigma) \ge 1 - \frac{1}{k^2}$
- Applying $k = 2$, we conclude that for any random variable X, there is at most $\frac{1}{4}$ chance that it is 2 SD or further away from its mean

Joint Distribution

Joint Probability Mass Function

- $f_{X,Y}(x, y) \ge 0,\ \forall (x, y) \in R_{X,Y}$
- $\sum_x \sum_y f_{X,Y}(x, y) = 1$
- $P((X, Y) \in A) = \sum_{(x,y) \in A} f_{X,Y}(x, y)$

Marginal Probability Mass Function

- $f_X(x) = \sum_y P(X = x, Y = y) = \sum_y f_{X,Y}(x, y)$
- $f_Y(y) = \sum_x P(X = x, Y = y) = \sum_x f_{X,Y}(x, y)$

Joint Probability Density Function

- $f_{X,Y}(x, y) \ge 0,\ \forall (x, y) \in R_{X,Y}$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dydx = 1$

Marginal Probability Density Function

- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy$
- $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx$

Conditional PDF/PMF

- $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$, provided $f_Y(y) > 0$

Properties:

1. For a fixed y, $f_{X|Y}(x|y) \ge 0$
2. **Discrete**: $\sum_x f_{X|Y}(x|y) = 1$
3. **Continuous**: $\int_{-\infty}^{\infty} f_{X|Y}(x|y)\,dx = 1$
4. For $f_X(x) > 0$, $f_{X,Y}(x, y) = f_{Y|X}(y|x)f_X(x)$

Independent Random Variables

X and Y are independent iff, $\forall x, y$,

- $f_{X,Y}(x, y) = f_X(x)f_Y(y)$
- $f_{X|Y}(x|y) = f_X(x)$

Expectation ($E[g(X, Y)]$)

- **Discrete**: $\sum_X \sum_Y g(x, y)f_{X,Y}(x, y)$
- **Continuous**: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)F_{X,Y}(x, y)\,dydx$

Properties:

1. $E(a_0 + a_1X_1 + \cdots + a_nX_n) = a_0 + a_1E(X_1) + \ldots$
2. **Discrete**: $E(XY) = \sum_{x,y} [xy\, f_{X,Y}(x, y)]$
3. **Cont**: $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f_{X,Y}(x, y)\,dydx$
4. X and Y independent $\implies E(XY) = E(X)E(Y)$

To solve for $E(X|Y = n)$:

1. Find $f_{X|Y}(x|n)$
2. Solve for $\sum_x xf_{X|Y}(x|n)$ **OR** $\int_{-\infty}^{\infty} xf_{X|Y}(x|n)$

Covariance ($\mathrm{cov}(X, Y)\,|\,\sigma_{X,Y}$)

- $\mathrm{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$

Properties:

1. $\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y)$
2. $\mathrm{cov}(X, X) = V(X)$ which is why $\sigma_{X,X} = \sigma_X^2$
3. $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$
4. $\mathrm{cov}(aX + b, cY + d) = ac \times \mathrm{cov}(X, Y)$
5. $V(aX + bY) = a^2V(X) + b^2V(Y) + 2ab \times \mathrm{cov}(X, Y)$
6. X, Y are independent $\implies \mathrm{cov}(X, Y) = 0$

Correlation Coefficient ($\rho_{X,Y}$)

- $\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$

Properties:

1. $-1 \le \rho_{X,Y} \le 1$
2. X, Y are independent $\implies \rho_{X,Y} = 0$
   ***Note**: converse is **not** true

Discrete Distribution

Discrete Uniform Distribution

If X assumes $x_1, x_2, \ldots, x_k$ with equal probability,

- $f_X(x) = P(X = x) = \begin{cases} \frac{1}{k}, & x = x_1, x_2, \ldots, x_k \\ 0, & \text{otherwise} \end{cases}$
- $E(X) = \sum xf_X(x) = \frac{1}{k}\sum_{i=1}^k x_i$
- $V(X) = E(X^2) - [E(X)]^2 = \sum (x - \mu)^2 f_X(x)$

Bernoulli Distribution ($X \sim \mathrm{Bern}(p)$)

- $f_X(x) = P(X = x) = p^x(1 - p)^{1-x}$, for $0 < p < 1$
- Probability distribution of a single experiment with only 2 outcomes (ie. $x = 0, 1$ only)

Properties:

1. $E(X) = p$ and $V(X) = p(1 - p)$

Binomial Distribution ($X \sim \mathrm{B}(n, p)$)

- $f_X(x) = \binom{n}{x}p^x(1 - p)^{n-x}$, for $0 < p < 1$
- Distribution of number of **successes** in n independent Bernoulli trials (ie. $n \in \mathbb{Z}^+$)

Properties:

1. $E(X) = np$ and $V(X) = np(1 - p)$
2. Only 2 possible outcomes: success or failure
3. p is constant and independent in each trial

Binomial Approximations

1. $n = 1 \implies \mathrm{B}(1, p) = \mathrm{Bern}(p)$
2. $(n \ge 20 \wedge p \le 0.05) \vee (n \ge 100 \wedge np \le 10) \implies \mathrm{B}(n, p) \approx \mathrm{Poisson}(np)$. If $p \to 1$, use $q = 1 - p$
3. $np > 5 \wedge n(1 - p) > 5 \implies \mathrm{B} \approx N(np, np(1 - p))$
   ***Note**: continuity correction:
   (a) $P(X = k) \approx P(k - \frac{1}{2} < X < k + \frac{1}{2})$
   (b) $P(a < X < b) \approx P(a + \frac{1}{2} < X < b - \frac{1}{2})$
   (c) $P(a \le X \le b) \approx P(a - \frac{1}{2} < X < b + \frac{1}{2})$

Geometric Distribution ($X \sim \mathrm{Geom}(p)$)

- $f_X(x) = (1 - p)^{x-1}p$, for $0 < p < 1$
- Distribution of number of trials required until first success is achieved (ie. $x = 1, 2, 3, \ldots$)
- X denotes number of trials till first success

Properties:

1. $E(X) = \frac{1}{p}$ and $V(X) = \frac{1-p}{p^2}$
2. $P(X > n) = (1 - p)^n$
3. $P(X > n + k|X > n) = P(X > k), \forall n, k \ge 1$

Negative Binomial Distribution ($X \sim \mathrm{NB}(r, p)$)

- $f_X(x) = \binom{x-1}{r-1}p^r(1 - p)^{x-r}$, for $0 < p < 1$
- Distribution of number of **trials** required in order to obtain r successes ($x = r, r + 1, \ldots$ and $r \in \mathbb{Z}^+$)

Properties:

1. $E(X) = \frac{r}{p}$ and $V(X) = \frac{(1-p)r}{p^2}$
2. $k = 1 \implies \mathrm{NB}(1, p) = \mathrm{Geom}(p)$

Poisson Distribution ($X \sim \mathrm{Poisson}(\lambda)$)

- $f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, for $x = 0, 1, 2, \ldots$
- Distribution within fixed interval of time or space

Properties:

1. $E(X) = \lambda$ and $V(X) = \lambda$
2. $E[X(X - 1)] = E(X^2) - E(X) = \lambda^2$
3. Number of successes in an interval are independent of those occurring in any other disjoint intervals

Continuous Distribution

Continuous Uniform Distribution ($X \sim U(a, b)$)

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

Properties:

1. $E(X) = \frac{a+b}{2}$ and $V(X) = \frac{(b-a)^2}{12}$
2. $P(c \le X \le d) = \int_c^d f_X(x)\,dx = \frac{d-c}{b-a}$

Exponential Distribution ($X \sim \text{Exp}(\lambda)$, $\lambda = \frac{1}{\mu}$)

- $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, \\ 0, \end{cases} \quad F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$

Properties:

1. $E(X) = SD(X) = \frac{1}{\lambda}$ and $V(X) = \frac{1}{\lambda^2}$
2. Memoryless: $P(X > s + t | X > s) = P(X > t)$
3. $P(c < X < d) = F_X(d) - F_X(c) = e^{-\lambda c} - e^{-\lambda d}$

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

- $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x, \mu < \infty$, $\sigma > 0$

Properties:

1. $E(X) = \mu$ and $V(X) = \sigma^2$
2. $\int_{\mu-\sigma}^{\mu+\sigma} f_X(x) \approx 0.68$; $\int_{\mu-2\sigma}^{\mu+2\sigma} f_X(x) \approx 0.95$;
3. μ affects center, σ^2 affects shape/spread
4. If $X_k \sim N(\mu_k, \sigma_k^2)$ and $W = \sum_{i=1}^n a_i X_i$, where $k \in \mathbb{N} \le n$. Then, $W \sim N(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$
   - LC of normal random var X is also normal

Standard Normal Distribution ($Z \sim N(0, 1)$)

- $\phi(x) = \frac{1}{\sqrt{s\pi}} e^{-\frac{x^2}{2}}$ • $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}}\, dy$
- $P(a < Y \sim N(\mu, \sigma^2) \le b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$

For $X \sim N(\mu, \sigma^2)$, to normalise for $P(X < c)$:

- $P(X < c) = P(\frac{X-\mu}{\sigma} < \frac{c-\mu}{\sigma}) = P(Z < \frac{c-\mu}{\sigma})$

Properties:

1. $-Z \sim N(0, 1)$ and $E(Z^i) = 0$, where $i \in \mathbb{Z}^+$ is odd
2. $P(Z < -x) = P(Z \ge x) = 1 - P(Z < x)$
3. $Y \sim N(\mu, \sigma^2) \implies \frac{Y-\mu}{\sigma} \sim N(0, 1)$
4. $X \sim N(0, 1) \implies aX + b \sim N(b, a^2), \forall a, b \in \mathbb{R}$

Sampling & Sampling Distributions

Random Sample

Let X be a random variable with probability distribution $f_X(x)$, and let $X_1, \dots, X_n$ be n independent random variables, then $(X_1, \dots, X_n)$ is a random sample of size n from population with distribution $f_X(x)$

Sampling Distribution of the Sample Mean ($\bar{X}$)

- $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ = sample mean of n random samples with population mean μ and SD σ
- Assume sampling from infinite population or a small fraction of a large finite population

Properties:

1. $\bar{X}$ is a random variable (since $X_1, \dots, X_n$ are too)
2. $E(\bar{X}) = E(X) = \mu$ and $V(\bar{X}) = \frac{V(X)}{n} = \frac{\sigma^2}{n}$
3. $P(|\bar{X} - \mu| > \varepsilon) \to 0$ as $n \to \infty$ (as sample size n increases, probability that sample mean $\bar{X}$ differs from population mean μ approaches 0)

Central Limit Theorem

- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, for n random samples, if either:
  1. Population is normally distributed OR
  2. Population not normally distributed but $n \ge 30$

Difference of Two Sample Means

Let $\bar{X}_1$, $\bar{X}_2$ represent sample mean of two **independent** random samples of size $n_1, n_2 \ge 30$ with mean μ_1, μ_2 and variance σ_1^2, σ_2^2 respectively

- $E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$
- $V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
- $(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ approximately

Sample Variance (S^2)

- $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{(\sum_{i=1}^n X_i^2) - n\bar{X}^2}{n-1}$
- $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$

Gamma Function ($\Gamma(\alpha)$)

- $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy$

Properties:

1. $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ via integration by parts
2. $\Gamma(1) = \int_0^\infty e^{-y}\, dy = 1$
3. $\Gamma(n) = (n-1)!$ for integral values of α (ie. $n \in \mathbb{Z}^+$)

Chi-square Distribution ($X \sim \chi^2(n)$)

- PDF of $\chi^2(n) = f_X(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2}\Gamma(n/2)}$, for $x \ge 0$

Properties:

1. $E(X) = n$, $V(X) = 2n$ (n = degrees of freedom)
2. $\chi^2(n) \sim N(n, 2n)$ approximately for large n
3. $X \sim N(\mu, \sigma^2) \implies (\frac{X-\mu}{\sigma})^2 \sim \chi^2(1)$
4. $\sum_{i=1}^n (\frac{X_i - \mu}{\sigma})^2 \sim \chi^2(n)$, if X_i are random samples
5. $\sum_{i=1}^k Y_i \sim \chi^2(\sum_{i=1}^k n_i)$, if $Y_i \sim \chi^2(n_i)$

Student t Distribution ($T \sim t(n)$)

- $T = \frac{Z}{\sqrt{U/n}} \sim t(n)$, if $Z \sim N(0, 1)$ and $U \sim \chi^2(n)$

Properties:

1. $E(T) = 0$ for $n > 1$ and $V(T) = \frac{n}{n-2}$ for $n > 2$
2. PDF is bell shaped and symmetrical at $x = 0$
3. $T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$, if a random sample of size n is independently drawn from a normal population
4. As $n \to \infty$ (≥ 30 in practice), $t(n) \approx N(0, 1)$

F Distribution ($F \sim F(n_1, n_2)$)

- $\frac{U/n_1}{V/n_2} \sim F(n_1, n_2)$, if $U \sim \chi^2(n_1)$ and $V \sim \chi^2(n_2)$

Properties:

1. $F \sim F(n, m) \implies \frac{1}{F} \sim F(m, n)$
2. $F_{n,m;1-\alpha} = \frac{1}{F_{m,n;\alpha}}$
3. $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$

Normal Distribution Estimation

- Statistic: not dependent on unknown params (θ)
- Estimator: uses some statistic to estimate θ
- $\hat{\Theta}$ is an unbiased estimator of $\theta \iff E(\hat{\Theta}) = \theta$ eg. $\bar{X}$ is unbiased estimator of μ as $E(\bar{X}) = \mu$
- Interval Estimate of θ: $\hat{\theta}_L < \theta < \hat{\theta}_U$, where $\hat{\theta}_L, \hat{\theta}_U$ depend on value/sampling distribution of statistic
- Confidence Interval (CI): an interval $(\hat{\Theta}_L, \hat{\Theta}_U)$ containing θ such that $P(\hat{\Theta}_L < \theta < \hat{\Theta}_U) = 1 - \alpha$

CI For μ With Known σ

Population is normal OR n is sufficiently large (≥ 30):

- $\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = (1 - \alpha)100\%$ CI
- $n \ge (z_{\alpha/2} \frac{\sigma}{e})^2$, where Margin of Error $e = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

CI For μ With Unknown σ

Population is normal AND n is small (< 30):

- $\bar{X} \pm t_{n-1;\alpha/2} \frac{S}{\sqrt{n}} = (1 - \alpha)100\%$ CI

Population is normal AND n is large (≥ 30):

- $\bar{X} \pm z_{\alpha/2} \frac{S}{\sqrt{n}} = (1 - \alpha)100\%$ CI

CI For $\mu_1 - \mu_2$ With Known $\sigma_1 \ne \sigma_2$

Populations are normal OR n_1, n_2 are large (≥ 30):

- $(\bar{X}_1 - \bar{X}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = (1 - \alpha)100\%$ CI

CI For $\mu_1 - \mu_2$ With Unknown $\sigma_1 \ne \sigma_2$

n_1, n_2 are large (≥ 30), replace σ_1, σ_2 with S_1, S_2:

- $(\bar{X}_1 - \bar{X}_2) \pm z_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = (1 - \alpha)100\%$ CI

CI For $\mu_1 - \mu_2$ With Unknown $\sigma_1 = \sigma_2$

Populations are normal AND n_1, n_2 are small (< 30):

- $(\bar{X}_1 - \bar{X}_2) \pm t_{n_1+n_2-2;\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, where $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$ = pooled variance

CI For $\mu_D = \mu_1 - \mu_2$ With Paired Data

Let $S_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1} = \frac{(\sum_{i=1}^n D_i^2) - n\bar{D}^2}{n-1}$, where $D_i = X_i - Y_i$, $\bar{D} = \frac{\sum_{i=1}^n D_i}{n}$ (use GC with D_i values).

Population is normal AND n is small (< 30):

- $\bar{D} \pm t_{n-1;\alpha/2} \frac{S_D}{\sqrt{n}} = (1 - \alpha)100\%$ CI

Population size n is large (≥ 30):

- $\bar{D} \pm z_{\alpha/2} \frac{S_D}{\sqrt{n}} = (1 - \alpha)100\%$ CI

CI For σ^2 With Known μ

Population is normal, regardless of n:

- $\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{n;\alpha/2}} < \sigma^2 < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{n;1-\alpha/2}}$

CI For σ^2 With Unknown μ

Population is normal, regardless of n:

- $\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}$

CI For σ_1^2/σ_2^2 With Unknown μ

Populations are normal, regardless of n:

- $\frac{S_1^2}{S_2^2} \frac{1}{F_{n_1-1,n_2-1;\alpha/2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} F_{n_2-1,n_1-1;\alpha/2}$

Hypothesis Testing

- Reject: hypothesis is false
- Accept: insufficient evidence to reject hypothesis
- Null Hypothesis, H_0: parameter to investigate
- Alternate Hypothesis, H_1: reject $H_0 \Rightarrow$ accept H_1
- Type I Error: rejecting H_0 when H_0 is true
- Type II Error: not rejecting H_0 when H_0 is false
- p-value: probability of obtaining test results at least as extreme as the results actually observed during the test, assuming null hypothesis is true

Level of Significance (α) & Power ($1 - \beta$)

- $\alpha = P(\text{Type I Error}) = P(\text{rejecting } H_0 | H_0)$
- $\beta = P(\text{Type II Error}) = P(\text{not rejecting } H_0 | \neg H_0)$
  1. Let $X_i \sim N(\mu_i, \sigma^2)$, where μ_i is original value from H_0, find $(1 - \alpha)100\%$ CI for $\bar{X}_i = (a, b)$
  2. Let $X_n \sim N(\mu_n, \sigma^2)$, where μ_n is a new value, find $P(a < X_n < b) = \beta$
- Power $= 1 - \beta = P(\text{rejecting } H_0 | \neg H_0)$

Steps For Hypothesis Testing

Let x be true param, and x_0 be hypothesised param, and let distribution to be used be D, thus $X \sim D$

1. Assume $H_0 : x = x_0$ is true, and decide H_1 and α
2. Calculate respective test statistic, θ
3. Find p-value using θ and respective $X \sim D$
4. p-value $< \alpha \implies$ reject H_0

H_1	p-value	Rejection Region
$x > x_0$	$P(X > \theta)$	$\theta > Q_{r;\alpha}$
$x < x_0$	$P(X < \theta)$	$\theta < Q_{l;\alpha}$
$x \ne x_0$	$2P(X > \|\theta\|)$	$\theta > Q_{r;\frac{\alpha}{2}}$ or $< Q_{l;\frac{\alpha}{2}}$

Hypothesis On μ With Known σ

Population is normal OR n is sufficiently large (≥ 30),

- Test Statistic: $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$

Hypothesis On μ With Unknown σ

Population is normal,

- Test Statistic: $T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$

Hypothesis On $\mu_1 - \mu_2$ With Known σ_1, σ_2

Population is normal OR n is sufficiently large (≥ 30),

- Test Statistic: $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$

Hypothesis On $\mu_1 - \mu_2$ With Unknown σ_1, σ_2

Population sizes are sufficiently large ($n_1, n_2 \ge 30$),

- Test Statistic: $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim N(0, 1)$

Populations are normal AND n_1, n_2 are small (< 30),

- Let $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$, and $n_p = n_1 + n_2 - 2$
- Test Statistic: $T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}} \sim t(n_p)$

Hypothesis On $\mu_1 - \mu_2$ With Paired Data

Population is normal AND n is small (< 30),

Let $\bar{D} = \bar{X} - \bar{Y}$, $\mu_D = \mu_1 - \mu_2$, $H_0 : \mu_D = \mu_{D,0}$,

- Test Statistic: $T = \frac{\bar{D} - \mu_{D,0}}{S_D/\sqrt{n}} \sim t(n-1)$

Population size is large ($n \ge 30$),

- Test Statistic: $Z = \frac{\bar{D} - \mu_{D,0}}{S_D/\sqrt{n}} \sim N(0, 1)$

Hypothesis On σ

Population is (approximately) normal,

- Test Statistic: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

Hypothesis On $\sigma_1 - \sigma_2$ With Unknown μ_1, μ_2

Populations are normal,

- Test Statistic: $F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$
- Two-tailed: p-value $= 2\,min\{P(F < \theta), P(F > \theta)\}$

Miscellaneous

Calculation	TI84 Command
$\bar{x}, \sum x, \sum x^2, S_x$	`stat` → `1-Var Stats`
$P(\text{B}(n, p) \le a)$ $P(\text{Geom}(p) \le a)$	`binomcdf(n,p,a)` `geometcdf(p,a)`

$P(a < N(0, 1) < b)$
$Q_{r;\alpha} = z_\alpha$
$Q_{l;\alpha} = -z_\alpha$
	`normalcdf(a,b,0,1)`
	`invNorm(α,0,1,RIGHT)`
	`invNorm(α,0,1,LEFt)`

$P(a < t(n) < b)$
$Q_{r;\alpha} = t_{n;\alpha}$
$Q_{r;\alpha} = -t_{n;\alpha}$
	`tcdf(a,b,n)`
	`invT(1-α,n)`
	`invT(α,n)`

$P(a < \chi^2(n) < b)$ $Q_{r;\alpha} = \chi^2_{n;\alpha}$ $Q_{l;\alpha} = \chi^2_{n;1-\alpha}$	χ^2`cdf(a,b,n)` use χ^2 table use χ^2 table
$P(a < F(n, m) < b)$ $Q_{r;\alpha} = F_{n,m;\alpha}$ $Q_{l;\alpha} = F_{n,m;1-\alpha}$	`Fcdf(a,b,n,m)` use F table use F table

- $Q_{r;\alpha}$ is the value where $P(\theta > Q_{r;\alpha}) = \alpha$
  - Area under graph **right** of $x = Q_{r;\alpha}$ is α
- $Q_{l;\alpha}$ is the value where $P(\theta < Q_{l;\alpha}) = \alpha$
  - Area under graph **left** of $x = Q_{l;\alpha}$ is α
- IBP: $\int_a^b u\, dv = [uv]_a^b - \int_a^b v\, du$ (LIATE for u)