MA1521 Cheat Sheet

by Howard Liu

AY2019/20 Semester 1

1 Basics of Probability

1-34 Operation of Events

Refer to CS1231S page 2 column 2 **"Set Properties"**.

1-107 Binomial Coefficient

1. $\binom{n}{r} = \binom{n}{n-r}$ for $r = 0, 1, \ldots, n$
2. $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$ for $1 \leq r \leq n$
3. $\binom{n}{r} = 0$ for $r < 0$ or $r > n$

1-128 Axioms of Probability

For each event A of the sample space S we assume that a number $Pr(A)$, which is called the **probability** of the event A, is defined and satisfies the following three axioms:

1. $0 \leq Pr(A) \leq 1$
2. $Pr(S) = 1$
3. if A_1, A_2, ... are **mutually exclusive** events (that is, $A_i \cap A_j = \emptyset$ when $i \neq j$), then

$$Pr(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} Pr(A_i)$$

1-141 Probability Properties

For any two events A and B,

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$

1-144 The Inclusion-Exclusion Principle

$Pr(\bigcup\limits_{i=1}^{n} A_i) = \sum_{i=1}^{n} Pr(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Pr(A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} Pr(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} Pr(\bigcap\limits_{i=1}^{n} A_i)$

1-194 Conditional Probability

The conditional probability of B given A, is defined as

$$Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)} \text{ , if } Pr(A) \neq 0$$

If events B_1, B_2 are **mutually exclusive** events, we have

$$Pr(B_1 \cup B_2|A) = Pr(B_1|A) + Pr(B_2|A)$$

1-216 Multiplication Rule

If we have 3 events A, B and C, we have

$$Pr(A \cap B \cap C) = Pr(A)Pr(B|A)Pr(C|A \cap B)$$

providing that $Pr(A \cap B) > 0$

1-243 Bayer's Theorem

$$Pr(A|B) = \frac{Pr(A)Pr(B|A)}{Pr(B)}$$

1-271 Independent Events

Two events A and B are independent iff.

$$Pr(A \cap B) = Pr(A)Pr(B)$$

1-273 Properties of Independent Events

1. $Pr(A|B) = Pr(A)$ and $Pr(B|A) = Pr(B)$
2. When two events (each with probability greater than 0) are **independent**, they cannot be **mutually exclusive**. Vice versa.
3. The sample space S and the empty set $\emptyset$ are independent of any events.
4. If $A \subset B$, then A and B are dependent unless $B = S$.
5. Properties of independence cannot be shown on a Venn diagram.
6. If A and B are independent, so are A and B', A' and B, A' and B'.

1-288 Pairwise and Mutually Independence

Mutually independence implies pairwise independence. However, the **reverse** does not hold - pairwise independence does not imply mutually independence.

2 Concepts of Random Variables

2-12 Equivalent Events

Two events A and B are equivalent iff. A **consists of all sample points, s, in S for which** $X(s) \in B$.

2-22 Probability Function

The probability of $X = x_i$ denoted by $f(x_i)$ must satisfy the following two conditions

1. $f(x_i) \geq 0$ for all x_i.
2. $\sum_{i=1}^{\infty} f(x_i) = 1$.

2-44 Probability Density Function

For any c and d s.t. $c < d$,

$$Pr(c \leq X \leq d) = \int_c^d f(x)dx$$

Note: for any specified value of X, say x_0, we have $Pr(X = x_0) = Pr(x_0 \leq X \leq x_0) = \int_{x_0}^{x_0} f(x)dx = 0$

2-60 Cumulative Distribution Function

For c.d.f, we have definition $F(x) = Pr(X \leq x)$.
If it is discrete random variable, then

$$F(x) = \sum_{t \leq x} f(t) = \sum_{t \leq x} Pr(X = t)$$

If it is continuous random variable, then

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Remark: $F(x)$ is non-decreasing.

2-64 Derive p.f and p.d.f from c.d.f

For a continuous random variable,

$$f(x) = \frac{dF(x)}{dx}$$

if the derivative exists.
Also, we have
$Pr(a \neq X \neq b) = Pr(a < X \neq b) = F(b) - F(a)$ for CRV.

2-87 Expected Values (Mean)

For DRV, we define the **mean or expected value** of X, denoted by $E(X)$ or μ_X as:

$$\mu_X = E(X) = \sum_x x f_X(x)$$

Remark: $E(X)$ does not necessary to be a value of X.
For CRV, it is defined as:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$$

Remark: Expectation of a RV exists provided the sum or integral exists.

2-103 Expectation of a Function of a RV

For DRV and CRV respectively,

1. $E[g(X)] = \sum_x g(x)f_X(x)$
2. $E[g(X)] = \int_{\infty}^{\infty} g(x)f_X(x)dx$

Special Case: $V(X)$ below and **k-th moment of X** which is $E(X^k)$

2-104 Variance

When $g(x) = (X - \mu_X)^2$, $E(g(x))$ is called the **variance** of X.

$$\sigma_X^2 = V(X) = E[(X - \mu_X)^2]$$

Remarks:

1. $V(X) \geq 0$
2. $V(X) = E(X^2) - [E(X)]^2$
3. Its principle square root is called **standard deviation**.

2-122 Properties of Expectation

When a and b are constants,

1. $E(aX + b) = aE(X) + b$
2. $V(aX + b) = a^2V(X)$

2-137 Chebyshev's Inequality

Let X be a random variable (DRV or CRV) with $E(X) = \mu$ and $V(X) = \sigma^2$, for any positive number k, we have:

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

Alternatively,

$$Pr(|X - \mu| \leq k\sigma) \leq 1 - \frac{1}{k^2}$$

Remarks: This applies for **all** distribution with finite mean and variance. Only a boundary is given and there is no guarantee that actual value is close to this boundary.

3 2D RV

3-10 Joint p.f./p.d.f for DRVs

1. $f_{X,Y}(x_i, y_i) \geq 0$
2. $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} f_{X,Y}(x_i, y_i) = 1$

3-21 Joint p.f./p.d.f for CRVs

1. $f_{X,Y}(x, y) \geq 0$ for all $(x, y) \in R_{X,Y}$
2. $\int\int_{(x,y)\in R_{X,Y}} f_{X,Y}(x, y)dxdy = 1$

3-30 Marginal Probability

For DRV:

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

For CRV:

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

3-41 Conditional Distribution of 2D RV

Conditional distribution of Y **fiven that** $X = x$ is given by:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} \text{ , if } f_X(x) > 0$$

Remarks: All requirements for 1D p.f/p.d.f still applies (see 2-22 Probability Function and 2-44 Probability Density Function above).

3-72 Uniformly Distributed

When we say X and Y are uniformly distributed over some area, it means that $f_{X,Y}$ is a constant within this boundary. We can let it be k, and use summation/integration in 3-10 or 3-21 to find this value.

3-84 Independent 2D RVs

Two RVs are said to be independent iff.

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) \text{ , for all } x, y$$

3-108 Expectation for 2D RVs

The expectation of $g(X, Y)$ is defined as

$$E[g(X, Y)] = \begin{cases} \sum_x \sum_y g(x, y)f_{X,Y}(x, y) \text{ , for DRVs} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f_{X,Y}(x, y)dxdy \text{ , for CRVs} \end{cases}$$

3-109 Covariance

When $g(X, Y) = (X - \mu_X)(Y - \mu_Y)$, it becomes the definition of **covariance** between two RVs.

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Remarks (a, b, c, d are constants):

1. $Cov(X, Y) = E(XY) - \mu_X\mu_Y$
2. If X and Y are **independent**, then $Cov(X, Y) = 0$. However, the reverse is not true.
3. $Cov(aX + b, cX + d) = acCov(X, Y)$
4. $V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCov(X, Y)$

3-112 Correlation Coefficient

The **Correlation Coefficient** of X and Y, denoted by $Cor(X, Y)$, $\rho_{X,Y}$ or ρ, is defined by:

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

Remarks:

1. $-1 \leq \rho \leq 1$
2. ρ is a measure of the **degree if linear relationship** between X and Y.
3. If X and Y are independent, then $\rho = 0$. However, the reverse is not true.

4 Special Probability Distributions

4-4 Discrete Uniform Distribution

All k random variables all have the same probability. Hence, we have the p.f.:

$$f_X(x) = \frac{1}{k} \text{ for } x = x_1, x_2, \ldots, x_k$$

and 0 otherwise.

Mean and variance:

- $\mu = \frac{1}{k}\sum_{i=1}^{k} x_i$
- $\sigma^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu)^2$ or $\sigma^2 = \frac{1}{k}(\sum_{i=1}^{k} x_i^2) - \mu^2$

4-10 Bernoulli Distribution

There are only two outcomes, 0 and 1. We have p.f.:

$$f_X(x) = p^x(1-p)^{1-x} \text{ , } x = 0, 1$$

where p is a parameters and $0 < p < 1$; 0 otherwise.
Mean and variance:

- $\mu = p$
- $\sigma^2 = p(1-p) = pq$

4-20 Binomial Distribution

For an RV X having a **binomial distribution**, it can be seen as the sum of n **independent Bernoulli trials**:

$$X = Y_1 + Y_2 + \cdots + Y_n \text{ , where } Y_i \text{ has Bern. Dist. with p}$$

Then we have its p.f.:

$$f_X(x) = \binom{n}{x} p^x(1-p)^{n-x} = \binom{n}{q} p^x q^{n-x}$$

Mean and variance:

- $\mu = np$
- $\sigma^2 = p(1-p) = npq$

4-39 Negative Binomial Distribution

NBD interests in the k-th success occurs on the x-th trial. We have p.f.:

$$Pr(X = x) = f_X(x) = \binom{x-1}{k-1} p^k q^{x-k}$$

For $x = k, k+1, k+2, \ldots$
Mean and variance:

- $\mu = \frac{k}{p}$
- $\sigma^2 = \frac{(1-p)k}{p^2}$

4-51 Poisson Distribution

This describes the number of success X occuring **during a given time interval or in a speficied region**.
Properties:

1. X in one time interval or region is **independent** of those in other **disjoint** time interval or region of space.
2. $f_X(x)$ during a **very short time or in a very small region** is proportional to the length of the time interval or the size of the region.
3. $f_X(x)$ for $x > 1$ is negligible in the condition of (2) above.

P.F.:

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ for } x = 0, 1, 2, 3, \ldots$$

where λ is the average number of successes occuring in the given time interval or specified region.
Mean and variance:

- $\mu = \lambda$
- $\sigma^2 = \lambda$

4-73 Poisson Approximation to Binomial D.

If $X \sim B(n,p)$, when $n \to \infty$ and $p \to 0$ in such a way that $\lambda = np$ remains constant as $n \to \infty$, then X will have a approximate Poisson distribution:

$$\lim_{p \to 0, n \to \infty} Pr(X = x) = \frac{e^{-np}(np)^x}{x!}$$

Remark: If p is close to 1, we interchange what we defined as success and failure to get a p close to 0.

4-81 Continuous Uniform Distribution

When $X \sim Uniform(a,b)$, its p.d.f. graph will be a rectangle with base a to b inclusive and height $1/(b-a)$.
Mean and variance:

- $\mu = \frac{a+b}{2}$
- $\sigma^2 = \frac{1}{12}(b-a)^2$

4-90 Exponential Distribution

P.D.F of X having this distribution:

$$f_X(x) = \alpha e^{-ax} \text{ for } x > 0$$

Note: $\int_{-\infty}^{\infty} f(x)dx = 1$
Mean and variance:

- $\mu = \frac{1}{\alpha}$
- $\sigma^2 = \frac{1}{\alpha^2}$

No Memory Property:

$$Pr(X > s + t | X > s) = Pr(X > t)$$

Example: Probability a bulb lasts for **next** 1 month after using it for 12 months is the same as that for it to last for the 1st 1 month as brand new.

4-105 Normal Distribution

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

Properties:

1. Its graph is bell-shaped and symmetrical about $x = \mu$
2. Maximum point occurs at $x = \mu$ with $f(x) = \frac{1}{\sqrt{2\pi}\sigma}$
3. The curve approaches the x axis asymptotically when going to either direction
4. Total area under the curve is 1
5. Two curves with same σ will have same shape (with different center if μ is different)
6. When σ increases, the curve flattens (reverse: the curve sharpens)

Mean and variance:

- $E(X) = \mu$
- $V(X) = \sigma^2$

where μ and σ^2 are the parameters of the distribution.

4-110 Std. Normal Distribution

If X has a normal distribution,

$$Z = \frac{X - \mu}{\sigma}$$

has a **standardized normal distribution**, where $E(Z) = 0$ and $V(Z) = 1$.

4-132 Normal Approx. to Binomial D.

If $X \sim B(n,p)$, we have $\mu = np$ and $\sigma^2 = np(1-p)$. Then as $n \to \infty$,

$$Z = \frac{X - np}{\sqrt{npq}} \text{ is approx. } \sim N(0,1)$$

4-136 Continuity Correction

This applies to normal approx. to $B(n,p)$. Known, $0 \leq X \leq n$:

1. $Pr(X = k) \approx Pr(k - \frac{1}{2} < X < k + \frac{1}{2})$
2. - $Pr(a \leq X \leq b) \approx Pr(a - \frac{1}{2} < X < b + \frac{1}{2})$
   - $Pr(a < X \leq b) \approx Pr(a + \frac{1}{2} < X < b + \frac{1}{2})$
   - $Pr(a \leq X < b) \approx Pr(a - \frac{1}{2} < X < b - \frac{1}{2})$
   - $Pr(a < X < b) \approx Pr(a + \frac{1}{2} < X < b - \frac{1}{2})$
3. $Pr(X \leq c) = Pr(0 \leq X \leq c) \approx Pr(-\frac{1}{2} < X < c + \frac{1}{2})$
4. $Pr(X > c) = Pr(c < X \leq n) \approx Pr(c + \frac{1}{2} < X < n + \frac{1}{2})$

5 Sampling and Sampling Distributions

5-12 Sampling from a Finite Population

From a population of size N, drawing n samples. If no replacement, each sample has probability off $\frac{1}{{}_N C_n}$ being chosen. If there is replacement, $\frac{1}{N^n}$.

5-31 Sampling Distribution of $\bar{X}$

For random samples of size n from infinite population or finite one with replacement having population mean μ and population standard deviation σ, sampling distribution of $\bar{X}$:

$$\mu_{\bar{X}} = \mu_X \text{ and } \sigma_{\bar{X}^2} = \frac{\sigma_X^2}{n}$$

That is,

$$E(\bar{X}) = E(X) \text{ and } V(\bar{X}) = \frac{V(X)}{n}$$

Law of Large Number: When population have a finite σ^2, as the sample size increases, probability that sample mean differs from population mean goes to 0.

5-37 Central Limit Theorem

The **sample distribution** of $\bar{X}$ is approximately normal with mean μ and variance σ^2/n if n is sufficiently large (say, ≥ 30).
If X is originally normally distributed, $\bar{X}$ is normally distributed regardless the size of n.

5-52 Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

If independent samples of sizes n_1 and n_2 (each ≥ 30) are drawn from two populations, with means μ_1, μ_2 and variances σ_1^2, σ_2^2,

$$\bar{X}_i - \bar{X}_2 \text{ approx. } \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

5-59 Chi-square Distribution

If Y is an RV and it has p.d.f.:

$$f_Y(y) = \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2} \text{ , for } y > 0$$

and 0 otherwise, Y is defined to have a **Chi-square distribution with n degrees of freedom**, denoted by $Y \sim \chi^2(n)$. n is a positive integer.
Mean and variance:

- $\mu = n$
- $\sigma^2 = 2n$

Summation: For independent $Y_1 \sim \chi^2(n_1), Y_2 \sim \chi^2(n_2), \ldots, \sum_{i=1}^k Y_i \sim \chi^2(\sum_{i=1}^k n_i)$.

5-62 Conversion to Chi-Square D.

1. If $X \sim N(0,1)$, then $X^2 = \chi^2(1)$.
2. If $X \sim N(\mu, \sigma^2)$, then $[(X-\mu)/\sigma]^2 \sim \chi^2(1)$.
3. When there are n random samples from a normal population, define

$$Y = \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}$$

then $Y \sim \chi^2(n)$.

5-67 Sample Variance

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

5-68 Sample Variance and Chi-sq.

If S^2 is from samples from a **normal** population having variance σ^2, then

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

5-69 The T-distribution

Suppose **independent** RVs $Z \sim N(0,1)$ and $U \sim \chi^2(n)$, and let

$$T = \frac{Z}{\sqrt{U/n}}$$

then the RV T follows **the t-distribution with n degrees of freedom**. That is,

$$T \sim t(n)$$

Properties:

1. The graph of t-distribution is symmetrical about y-axis, and is very close to that of the standard normal distribution.
2. $\lim_{n\to\infty} f_T(t) = f_Z(t)$

Mean and variance:

- $E(T) = 0$
- $V(T) = n/(n-2)$ for $n > 2$

5-80 The F-distribution

Let $U \sim \chi^2(n_1)$ and $V \sim \chi^2(n_2)$, then

$$F = \frac{U/n_1}{V/n_2}$$

is called a F distribution with (n_1, n_2) degrees of freedom.
Mean and variance:

- $E(X) = n_2/(n_2 - 2)$ with $n_2 > 2$
- $V(X) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$ for $n_2 > 4$

Remark: If $F \sim F(n,m)$, then $1/F$ $F(m,n)$.

6 Estimation Based on Normal Distribution

6-11 Interval Estimation of μ

Suppose σ^2 is known. Let

$$\hat{\Theta}_L = \bar{X} - 2\frac{\sigma}{\sqrt{n}} \text{ and } \hat{\Theta}_U = \bar{X} + 2\frac{\sigma}{\sqrt{n}}$$

then we have an interval estimator of μ:

$$(\bar{X} - 2\frac{\sigma}{\sqrt{n}}, \bar{X} + 2\frac{\sigma}{\sqrt{n}})$$

6-12 Unbiased Estimator

A statictic $\hat{\Theta}$ is said to be an **unbiased estimator** of the parameter θ if:

$$E(\hat{\Theta}) = \theta$$

Examples: $\bar{X}$ is an unbiased estimator of μ and S^2 is an unbiased estimator of σ^2.

6-17 Interval Estimation

We seek a random interval $(\hat{\Theta}_L, \hat{\Theta}_U)$ containing θ with a given probability $1-\alpha$. That is,

$$Pr(\hat{\Theta}_L < \theta < \hat{\Theta}_U) = 1 - \alpha$$

and this is called **a** $(1-\alpha)100\%$ **confidence interval** for θ.

6-22 Known Variance Case

When population

1. has known variance and,
2. is normal or n is sufficiently large (CLT)

, we can have the interval given by

$$(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$$

and the size of error can be given by

$$Pr(\left|\bar{X} - \mu\right| < z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

6-27 Margin of Error

Let e denote the **margin of error**. We want:

$$Pr(\left|\bar{X} - \mu\right| \leq e) \geq 1 - \alpha$$

For a given margin of error e, the sample size is given by

$$n \geq (z_{\alpha/2}\frac{\sigma}{e})$$

6-32 Unknown Variance Case

It needs to satisfy:

1. unknown population variance
2. the population is **normal or very closed to normal**
3. the sample size is **small**

, then we let

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Hence,

$$Pr(-t_{n-1;\alpha/2} < T < t_{n-1;\alpha/2}) = 1 - \alpha$$

Or:

$$Pr(\bar{X} - t_{n-1;\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{n-1;\alpha/2}\frac{S}{\sqrt{n}}) = 1 - \alpha$$

6-36 Unknown Variance with Large n

When n is large, we simply replace *sigma* with S in the section 6-22 Known Variance Case above.

6-43 CI for independent $\bar{X}_1 - \bar{X}_2$

X_1 and X_2 have to be independent. We simply replace:

- $\bar{X}$ with $\bar{X}_1 - \bar{X}_2$
- μ with $\mu_1 - \mu_2$
- $\frac{\sigma}{\sqrt{n}}$ with $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

in the section 6-22 Known Variance Case and 6-27 Margin of Error above.

6-56 Unknown but Equal Variance

Conditions are the same as 6-32 Unknown Variance Case above. Let $\sigma_1 = \sigma_2 = \sigma$. Then, σ^2 can be estimated by the pooled variance:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Then, substituting S_p^2 for σ^2, we get this statistic:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim t_{n_1+n_2-2}$$

Hence, let

$$d = t_{n_1+n_2-2;\alpha/2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

We have the confidence interval be:

$$(\bar{X}_1 - \bar{X}_2) - d < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + d$$

6-64 Unknown but Equal Variance with Large n

Replace $t_{n_1+n_2-2;\alpha/2}$ by $z_{\alpha/2}$ in the above formula.

6-70 CI for paired data

Observations in two samples made from the **same individual** are related and hence form a pair. Consider $d_i = x_i - y_i$ of paired observations. We assume d is normal and we then have:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

and point estimate of σ_D^2:

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

Hence we have the CI for d:

$$\bar{d} - t_{n-1;\alpha/2}(\frac{S_D}{\sqrt{n}}) < \mu_D < \bar{d} + t_{n-1;\alpha/2}(\frac{S_D}{\sqrt{n}})$$

6-73 CI for paired data with large n

For **sufficiently large** sample, we can replace $t_{n-1;\alpha/2}$ by $z_{\alpha/2}$ above.

6-78 CI for variances

Consider $X_1, X_2, \ldots, X_n$ from (approximate) $N(\mu, \sigma^2)$ distribution. Then a point estimator of σ^2:

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

We have CI when μ is known:

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

When μ is unknown:

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

Remark: When we want to find CI for σ, we just square root both sides in the above inequalities.

6-90 CI for Ratio of Two Variances

Consider samples $X_1, X_2, \ldots, X_{n_1}$ from a $N(\mu_1, \sigma_1^2)$ population, and samples $Y_1, Y_2, \ldots, Y_{n_2}$ from a $N(\mu_2, \sigma_2^2)$ population, and μ_1, μ_2 are **unknown**. Then we have the CI:

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