ST2334 – Probability and Statistics AY24/25, Y2S2 Notes

Sim Ray En Ryan

29 April 2025

Contents

1 Basic Concepts of Probability

1.1 Definitions

A statistical experiment (such as rolling dice) produces data and observations.

1.1.1 Sample Space S

S is the set of all possible outcomes, such as $S = \{1, 2, 3, 4, 5, 6\}$.

1.1.2 Event

An event is a subset of the sample space:

- Dice is a natural number $= \{1, 2, 3, 4, 5, 6\} = S$ // This is a sure event

- Dice is odd $= \{1, 3, 5\}$

- Dice is 7 $= \{\}$ // This is a null event

1.2 Operations and Relations

- Disjunction / Union $\cup$

- Conjunction / Intersection $\cap$

- Negation / Complement $\neg$

- Containment / Subset $\subset$

    - In ST2334, $\subset \equiv \subseteq$

- Equivalence $=$

    - $A = B \implies A \subset B \wedge B \subset A$

- Mutually Exclusive $A \wedge B = \{\}$

- Independence $\perp$ $P(A \wedge B) = P(A)P(B)$

1.3 Counting Methods

1.3.1 Permutations

When order matters.
$$nPr(n, r) = \frac{n!}{(n-r)!} \tag{1}$$

4

1.3.2 Combinations

When order does not matter.

$$nCr(n, r) = \binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{2}$$

1.4 Axioms of Probability

1.4.1 Axioms

$$\forall A, 0 \le P(A) \le 1 \tag{3}$$

$$P(S) = 1 \tag{4}$$

$$P(A \wedge B) = 0 \implies P(A \vee B) = P(A) + P(B) \tag{5}$$

1.4.2 Other Properties

$$P(0) = 0 \tag{6}$$

$$P(\bigwedge_{i=1}^{n} E_i) = 0 \implies P(\bigvee_{i=1}^{n} E_i) = \bigvee_{i=1}^{n} P(E_i) \tag{7}$$

$$P(A') = 1 - P(A) \tag{8}$$

$$A \subset B \implies P(A) \le P(B) \tag{9}$$

$$A \subset B \implies A \cup B = B \tag{10}$$

$$A \cup B = A \cup (B \cap A') \tag{11}$$

$$A \cap B \cap A' = 0 \tag{12}$$

1.5 Conditional Probability

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{13}$$

Multiplication Rule: $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B), P(A) \neq 0 \wedge P(B) \neq 0$ $$\tag{14}$$

Inverse Probability Formula: $P(A|B) = \dfrac{P(A)P(B|A)}{P(B)}$ $$\tag{15}$$

$$P(B) = P(A)P(B|A) + P(A')P(B|A') \tag{16}$$

Bayes's Theorem: $P(A \mid B) = \dfrac{P(B \mid A)P(A)}{\sum_i P(B \mid A_i)P(A_i)}$ $$\tag{17}$$

2 Random Variables

2.1 Definitions

2.1.1 Random Variable

A function X which assigns a real number to every $s \in S$ is called a random variable.

$$\mathbf{X} : S \mapsto \mathbb{R}. \tag{18}$$

2.1.2 Range Space

The range space of X is the set of all possible values.

$$R_x = \{x | x = X(s), s \in S\} \tag{19}$$

If X = number of heads, and 2 coins are flipped, $R_x = \{0, 1, 2\}$. This must be a subset of S. Upper case letters denote random variables, and lower case letters denote observed values.

2.2 Probability Distribution

2.2.1 Discrete

The number of values in R_X is finite or countable.

Probability [mass] function: $f(x) = \begin{cases} P(X = x), & \text{for } x \in R_X; \\ 0, & \text{for } x \notin R_X. \end{cases}$ $$\tag{20}$$

$$\forall x_i \in R_x, f(x_i) \geq 0 \tag{21}$$

$$\forall x_i \notin R_x, f(x_i) = 0 \tag{22}$$

$$\sum_{i=1}^{\infty} f(x_i) = 1 \tag{23}$$

2.2.2 Continuous

The number of values in R_x is an interval or a collection of intervals

$$\forall x \in R_x, f(x) \geq 0, \forall y \notin R_x, f(y) = 0 \tag{24}$$

$$\int_{R_x} f(x)dx = 1 \tag{25}$$

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx. \tag{26}$$

$$P(X = x_0) = \int_{x_0}^{x_0} f(x)\, dx = 0. \tag{27}$$

2.3 Cumulative Distribution Function, cdf

The following applies for both discrete and continuous random variables. In both cases, F(x) is non-decreasing.

$$F(x) = P(X \leq x) \tag{28}$$

2.3.1 Discrete Variables

$$F(x) = \sum_{t \in R_X, t \leq x} f(t) = \sum_{t \in R_X, t \leq x} P(X = t) \tag{29}$$

2.3.2 Continuous Variables

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \tag{30}$$

$$f(x) = \frac{dF(x)}{dx}. \tag{31}$$

2.4   Expectation and Variance

2.4.1   Expectation

The expected value $E(X)$ is the average value of X, if the experiment is repeated many times.

$$\text{Discrete: } \mu_x = E(X) = \sum_{x_i \in R_X} x_i f(x_i) \tag{32}$$

$$\text{Continuous: } \mu_X = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{x \in R_X} x f(x)\, dx. \tag{33}$$

For continuous variables, it is not necessarily a possible value of random variable X.

- $E(aX + b) = aE(X) + b$

- $E(X + Y) = E(X) + E(Y)$

2.4.2   Variance

$$\sigma_X^2 = V(X) = E\big((X - \mu_X)^2\big). \tag{34}$$

- $V(X) \geq 0$ for any X. $V(X) = 0 \iff P(X = E(X)) = 1$

- $V(aX + b) = a^2 V(X)$

- $\sigma_x = \sqrt{V(X)}$

3   Joint Distributions

3.1   Multiple Random Variables

(x, y) is a two dimensional random vector/variable. Can be increased to become an n-dimension random vector/variable. This can be discrete or continuous depending on how many possible outcomes there are. In ST2334, (X, Y) is only discrete/continuous if both X and Y are discrete/continuous.

$$R_{x,y} = \{(x, y) | x = X(s), y = Y(s), s \in S\} \tag{35}$$

3.1.1   Joint Probability Function

$$\forall (x,y) \in R_{x,y}, f_{x,y}(x,y) = P(X=x, Y=y) \tag{36}$$

$$\forall (x,y) \in_{x,y}, f_{x,y}(x,y) \geq 0 \tag{37}$$

$$\forall (x,y) \notin_{x,y}, f_{x,y}(x,y) = 0 \tag{38}$$

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P(X=x_i, Y=y_j) = \sum \sum_{(x,y) \in R_{X,Y}} f(x,y) = 1 \tag{39}$$

$$P((X,Y) \in A) = \sum \sum_{(x,y) \in A} f_{X,Y}(x,y). \tag{40}$$

$$P((X,Y) \in D) = \iint_{(x,y) \in D} f_{X,Y}(x,y)\, dy\, dx \tag{41}$$

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x,y)\, dy\, dx \tag{42}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = \iint_{(x,y) \in R_{X,Y}} f_{X,Y}(x,y)\, dx\, dy = 1. \tag{43}$$

3.2   Marginal and Conditional Distributions

$$f_X(x) = \sum_y f_{x,y}(x,y) = \int_{-\infty}^{\infty} f_{x,y}(x,y) dy \tag{44}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \tag{45}$$

3.3   Independent Random Variables

X and Y are independent $\iff f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$
$$\tag{46}$$

For independence, $R_{x,y}$ needs to form the product space, and can be factorized into the product of two functions and a constant.

3.4 Expectation and Covariance

$$E(g(X,Y)) = \sum_x \sum_y g(x,y) f_{X,Y}(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) \, dy \, dx. \tag{47}$$

$$\text{cov}(X,Y) = E[(X - E(X))(Y - E(Y))] \tag{48}$$

$$= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) \tag{49}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) \, dx \, dy. \tag{50}$$

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y) \tag{51}$$

$$\text{X and Y are independent} \implies \text{cov}(X,Y) = 0 \text{ (Note the converse is not true)} \tag{52}$$

$$\text{cov}(aX + b, cY + d) = ac \cdot \text{cov}(X,Y). \tag{53}$$

$$V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab \cdot \text{cov}(X,Y). \tag{54}$$

$$V(X_1 \pm X_2 \pm \cdots \pm X_n) = V(X_1) + V(X_2) + \cdots + V(X_n). \tag{55}$$

$$V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n) + 2\sum_{j > i} \text{cov}(X_i, X_j). \tag{56}$$

4 Special Probability Distributions

4.1 Discrete Distributions

For discrete distributions, the range space R_x is always finite or countable.

4.1.1 Discrete Uniform Distributions

Every one of the k elements has an equal $\frac{1}{k}$ probability.

4.1.2 Bernoulli Trials

When an observation can be a success or failure. X is the number of successes in a trial, and the variable is known as a Bernoulli random variable. p is the probability of success, and the PMF is p if x is 1, and 1 - p if x is 0. The mean is p, and the variance is p(1 - p).

A Bernoulli process consists of repeatedly performing identical and independent Bernoulli trials.

4.1.3 Binomial Distribution

The probability of getting exactly x successes in n trials. The expected value is np and the variance is np(1-p).

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{for } x = 0, 1, 2, \ldots, n.$$

4.1.4 Negative Binomial Distribution

The expected number of trials, or probability that n trials need to occur, before x successes occur.

$$f_X(x) = P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad \text{for } x = k, k+1, k+2, \ldots$$

For GDC, do binomPdf(number of trials - 1, chance, number of success - 1) then multiply by chance. The expected value of X is k/p and the variance is (1 - p)k / pp.

If A has a 0.55 chance of winning each game, and the first team to 4 games out of 7 wins, then

X = number of game needed to win:

P(X = 6) = binomPdf(5, 0.55, 3) * 0.55

Win = P(4 ¡= X ¡= 7) = Sum 3 6 binomPdf(i, 0.55, 3) * 0.55.

4.1.5 Geometric Distribution

Same as negative but only for the first success to occur. The expected value is 1/p and variance is 1-p/pp as per above, but k = 1.

$$f_X(x) = P(X = x) = (1-p)^{x-1} p.$$

GDC calculation is same as before, where number of success - 1 = 0.

4.1.6 Poisson Process

When some of the arguments are limited. λ ¿ 0 is the expected number of occurrences in a given number of trials. Thus, E(X) is λ and Variance is also λ. If the poisson process occurs over multiple trials, just multiple λ by the number of trials, since if it expected to happen λ times in 1 trials, it is then expected to happen 2λ times in 2 trials.

Poisson Approximation to Binomial In binomial distribution, if the number of trials is infinite and if the chance of success is 0 (such that $\lambda = np$ is roughly constant, then the functions approximate each other. In particular, $n \geq 20, p \leq 0.05$, or $n \geq 100, np \leq 10$.

4.2 Continuous Distributions

For uncountable ranges.

4.2.1 Continuous Uniform Distributions

For every $a \leq x \leq b$, $f_x(x) = \frac{1}{b-a}$. The expected value is the average of a and b, and the variance is $\frac{(b-a)^2}{12}$ 12 appears because it is 3 * 4. The cumulative distribution is also as follows:

$$F_X(x) = \begin{cases} 0, & \text{for } x < a; \\ \frac{x-a}{b-a}, & \text{for } a \leq x \leq b; \\ 1, & \text{for } x > b. \end{cases}$$

4.2.2 Exponential Distribution

The continuous version of geometric distribution. Lambda is a rate parameter and represents the number of events that occur per unit time.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

For the cumulative case:

$F_X(x) = 1 - e^{-\lambda x}$. For x = 0, Fx(x) = 0.

The expected value is 1 / lambda, and the variance is 1 / lambda lambda. It is memoryless, in the sense that if s time has already passed then the probability that it will last another t time is the same as if the bulb was new and passes t time.

4.2.3 Normal Distribution

Has parameters of mean (expected value) and variance (sigma sq). It is positive over the real line, symmetrical about x = mu, and bell-shaped. Two curves are identical with the same variance, only differing by their x-position.

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

The peak (at the symmetrical line) has y = 1 / sqrt 2 pi. As sigma increases, the curve flattens. Convert a function to its standard form to calculate.

4.2.4 Quantiles

$z_{0.05} = 1.645, z_{0.01} = 2.326$

4.2.5 Continuity Correction

Normal also simulates binomial when n is large and p is constant. In particular, when np ¿ 5 and n(1-p) ¿ 5. When simulation, it is common to add or subtract half depending on the bound.

5 Sampling and Sampling Distributions

5.1 Population and Sample

Populations can be finite or infinite, which may also depend if the observations are numerical or categorical. A sample is any subset of a population. Some finite populations are better treated as infinite populations as it is impossible to observe all its values.

5.1.1 Populations

5.2 Random Sampling

5.2.1 Simple Random Sample of Finite Population

Everyone has the same chance of inclusion, representative if the population, and has low chance that sample is seriously biased in some way. For a population size of N and sample size of n, there is N choose n possible samples.

5.2.2 Simple Random Sample of Infinite Population

Sampling is random if with each draw, every element has the same probability of being selected, and successive draws are independent. SRS on a finite population with replacement is equivalent to and infinite population.

5.3 Sampling Distribution of Sample Mean

The mean (X bar) is the sum of all X divided by the number of X. The variance:
$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

5.3.1 Statistic

A function of observations taken is known as a statistic. For the above example, a realization of the statistic X bar is x bar, which contains the actual values of x. Similar for variance. Since it is just a statistic, the realization will vary from sample to sample.

5.3.2 Sampling Distribution

Mean (centre) and variance (spread) of X bar. The expectation of X bar is equal to the population mean. As the population gets larger and larger, the variance of expectation of X bar gets smaller and smaller. The law of large numbers states that:
$$P(|\bar{X} - \mu| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty.$$

5.3.3 Standard Error

$\sigma_{\bar{X}}$ different from σ which is standard deviation of the points itself, rather than the means of the samples taken.

5.4 Central Limit Theorem

If a sample size has mean μ and finite variance σ^2, as n approaches infinity, then X bar approaches a normal distribution of mean μ and variance of $\frac{\sigma^2}{n}$. In other words, with a large enough sample size of independent samples, distribution is normal. If the sample is already normal, then the normal distribution fit exactly.

5.5 Other Sampling Distributions

5.5.1 χ^2 Distributions

Normal distributions with n degrees of freedom. A χ^2 distribution is a family of curves, and have a long right tail. As n increases, it resembles a normal distribution more N(n, 2n). Adding to Chi sq together adds their degrees of freedom together.

5.5.2 Student's t Distributions

Z: N(0, 1), and U: Chi sq (n). If Z and U are independent,
$T = \frac{Z}{\sqrt{U/n}}$ with n degrees of freedom. Approaches N(0,1) as n increases. The expected value is 0 and variance is n / (n-2) for n ¿ 2.

If X is independent and identically distributed normal with mean mu and variance sigma sq, $\frac{Xbar - \mu}{S/\sqrt{n}}$ follows a t distribution with n-1 degrees of freedom/

5.5.3 F Distributions

For two chi sq distributions U: X2(m) and V: X2(n),
$F = \frac{U/m}{V/n}$ follows F distribution with (m, n) degrees of freedom.

6 Estimation

6.1 Point Estimation

When a single number is calculated to estimate a population parameter. The formula that describes this is a point estimator, and the resulting number is the point estimate.

If two numbers are used to describe a range, then it is an interval estimate.

For example, the point estimator would be

$(Xbar = X_1 + X_2 + X_3 + X_4)/4$

while the point estimate itself would be

$xbar = (6 + 1 + 3 + 4)/4 = 5$

6.1.1 Unbiased estimators

A good estimator should be unbiased. If the estimated value of X is the value of x, then it is an unbiased estimator of x. For example, for a uniform distribution U(0, x), then X bar

14

is not an unbiased estimator of x: 2X bar is. Also note that randomly picking just one X is also a valid unbiased estimator.

6.1.2   Maximum error of estimate

Although unbiased, X bar is usually not the average, and X bar - average measures the difference in estimation. With large numbers, things follow a normal distribution. Then, with a probability of 1 - a, there is a maximum error of estimate of $z_{a/2}\frac{\sigma}{\sqrt{n}}$. This z value is usually obtained from software.

Table 1: Different Cases

Case	Population	σ	n	Statistic	E	n for desired E_0 and α
I	Normal	known	any	$Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$	$\left(\dfrac{z_{\alpha/2} \cdot \sigma}{E_0}\right)^2$
II	any	known	large	$Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$	$\left(\dfrac{z_{\alpha/2} \cdot \sigma}{E_0}\right)^2$
III	Normal	unknown	small	$T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$	$t_{n-1;\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$	$\left(\dfrac{t_{n-1;\alpha/2} \cdot s}{E_0}\right)^2$
IV	any	unknown	large	$Z = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$	$z_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$	$\left(\dfrac{z_{\alpha/2} \cdot s}{E_0}\right)^2$

6.2   Confidence Intervals for the Mean

Slightly more useful as the point estimator is almost never correct. Instead, be 1 - alpha certain that it lies between a and b. (a, b) is then called the (1 - alpha) confidence interval. To calculate the interval. If you are super unlucky, the mean is not within this interval. Once an interval is constructed, the mean is either in it or not. Since the mean is usually never known, there is no way to determine, and the only other way is to keep repeating the procedure. About (1-a) of the confidence intervals will contain the mean.

Table 2: Confidence Intervals for Different Cases

Case	Population	σ	n	Confidence Interval
I	Normal	known	any	$\bar{x} \pm z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$
II	any	known	large	$\bar{x} \pm z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$
III	Normal	unknown	small	$\bar{x} \pm t_{n-1;\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$
IV	any	unknown	large	$\bar{x} \pm z_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$

6.3 Comparing Two Populations

When comparing the statistic of two populations, it can be independent samples (complete randomization) where all observations must be independent, or matched pair samples (randomization between matched pairs: number of samples from both population are the same), where within the pair, observations may be dependant, but between pairs is independant.

6.4 Independent Samples: Unequal Variances

When trying to find $\delta = \mu_1 - \mu_2$, the difference in a statistic of two populations, then the confidence interval is $\left(\bar{X} - \bar{Y}\right) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ Both populations must be normal or large enough. Replace with sample variance if needed.

6.5 Independent Samples: Equal Variances

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

$\left(\bar{X} - \bar{Y}\right) \pm t_{n_1+n_2-2;\alpha/2} \cdot S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

(It is p much just taking the sigma out). Can assume that variance is equal when the ratio is between 0.5 and 2. This is because statistic is not sensitive to small differences in population variance.

6.6 Paired Data

Must take the difference between each pair, then operate on that new set of data.

$\bar{d} \pm z_{\alpha/2} \cdot \frac{s_D}{\sqrt{n}}$

7 Hypothesis Tests

7.1 Hypothesis Tests

7.1.1 Null and Alternate Hypothesis

The null hypothesis is what you believe it is, the default assumption. The hypothesis we are trying to prove is the alternate hypothesis. Usually, the alternate hypothesis states that the null hypothesis is wrong, and can be rejected in favour of the alternate. Otherwise, it fails to reject (not accepts) the null hypothesis.

Sided tests If the null hypothesis is that mu = 100, and the alternate hypothesis is that mu ¿ 100, then it is one sided. If the alternate hypothesis is mu != 100, then it is two sided.

7.1.2   Level of Significance

There are two outcomes:

- Reject H0 and conclude H1, or

- Do not reject H0 and conclude H0

If H0 is true but H0 is rejected, it is a Type 1 error. If H0 is false but it is not rejected, then it is a Type 2 error. The chance of making a Type 1 error is the chance alpha = H0 is rejected given H0 is true. The chance of making a Type 2 error is beta = the chance H0 is not rejected given H0 is false. The level of significance of the test is alpha, and the power of the test is 1 - beta = chance H0 is rejected given H0 is false.

7.2   Hypothesis Concerning the Mean

7.2.1   Known Variance

Hypothesis Test for the Mean: Known variance
Consider the case where

- the population variance σ^2 is known; AND

- the underlying distribution is normal; OR

- n is sufficiently large (say, $n \geq 30$).

For the null hypothesis $H_0 : \mu = \mu_0$, the test statistic is given by

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

Let z be the observed Z value. For the alternative hypothesis

- $H_1 : \mu \neq \mu_0$, the rejection region is

$$z < -z_{\alpha/2} \quad \text{or} \quad z > z_{\alpha/2}.$$

- $H_1 : \mu < \mu_0$, the rejection region is

$$z < -z_\alpha.$$

- $H_1 : \mu > \mu_0$, the rejection region is

$$z > z_\alpha.$$

7.2.2 p-Values

If p value ¡ a, reject H0. Else, do not reject H0.

Hypothesis Test for the Mean: Unknown variance

Consider the case where

- the population variance σ^2 is unknown; AND

- the underlying distribution is normal.

For the null hypothesis $H_0 : \mu = \mu_0$, the test statistic is given by

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}.$$

Let t be the observed T value. For the alternative hypothesis

- $H_1 : \mu \neq \mu_0$, the rejection region is

$$t < -t_{n-1;\alpha/2} \quad \text{or} \quad t > t_{n-1;\alpha/2}.$$

- $H_1 : \mu < \mu_0$, the rejection region is

$$t < -t_{n-1;\alpha}.$$

- $H_1 : \mu > \mu_0$, the rejection region is

$$t > t_{n-1;\alpha}.$$

7.3 Tests Comparing Means: Independent Samples

$Z = \frac{(\bar{X}-\bar{Y})-\delta_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim N(0,1).$

Table 3: Hypothesis Test for Difference of Means

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

7.4 Tests Comparing Means: Paired Data

$T = \frac{\bar{D}-\mu_{D_0}}{S_D/\sqrt{n}}.$