Chapter 1 Stuff

Construct a **stem-and leaf plots**:

1. Split each observation into a

• Stem: one or more of the leading, or left-hand, digits; and a

• Leaf: the trailing, or remaining, digit(s) to the right.

2. Write the stems in a column, from the smallest to the largest. Include all stems between the smallest and largest, even if there are no corresponding leaves.

3. List all the digits of each leaf next to its corresponding stem. It is not necessary to put the leaves in increasing order, but make sure the leaves line up vertically.

4. Indicate the unites for the stems and leaves.

Example: Number of weekly client of one store are recorded. Construct the stem-and-leaf plot, and describe the distribution. $n=16$

420 395 208 581 443 353 496 528 544 389 399 498 627 618 808 469

1. split the data

42 | 0 too many stems

4 | 20

stem, leaf

2. 208 - 808

3.

Stem	Leaf
2	08
3	95 53 89 99
4	20 43 96 98 69
5	81 28 44
6	27 18
7	
8	08

← Peak

Stem: hundreds of digits
Leaf: tens and ones digits

Distribution
1. Representative value: 300-400
2. Spread: not too far from the center
3. Gap: One minor gap at 700
4. Extent of symmetry: not symmetric
5. Peak: one peak at 400
6. Outlier: example value would be 1600

Histogram

→ Equal Class Width
# of rows/classes: $\sqrt{n}$
Width of a class: max - min

→ Unequal Class Width:
# of rows/classes: anything
Y-axis: density: $\frac{\text{rel freq}}{\text{width}}$

Mean $(\bar{x})$

Median $(\tilde{x})$

Trimmed Mean $(\bar{x}_{tr}(p))$

↳ $n \cdot p$ = amount of numbers to trim

Sample Proportion of Success $(\hat{p})$

$$\hat{p} = \frac{n(S)}{n}$$

Sample Range

↳ $R = x_{max} - x_{min}$

Sample Variance (s^2)

↳ $$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Sample Standard Deviation (s)

$$s = \sqrt{s^2}$$

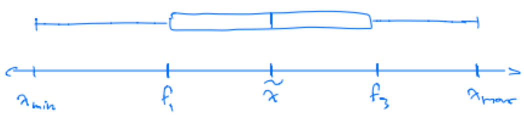
Lower Forth (f_1)

Upper Forth (f_3)

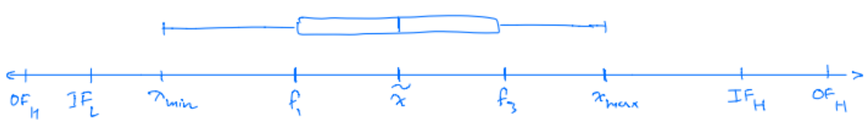
Fourth Spread (f_s)

↳ $f_s = f_3 - f_1$

Simple Box Plot



Modified Box Plot



Inner Fence $= [f_1 - 1.5f_s \,,\; f_3 + 1.5f_s]$

Outer Fence $= [f_1 - 3f_s \,,\; f_3 + 3f_s]$

Example 1.21 Sodium content values in food product. Construct simple box-plot, boxplot with outliers using the sample.

$n=12$

211 408 171 178 359 249 205 203 201 223 234 256.

mild

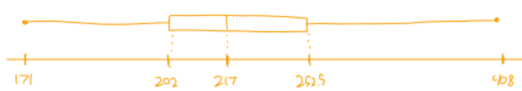
Extreme

171 178 201 | 203 205 211 | 223 234 249 | 256 359 408

$\tilde{x} = \frac{211+223}{2} = 217$ $\quad f_1 = \frac{201+203}{2} = 202$ $\quad f_3 = \frac{249+256}{2} = 252.5$

$x_{min} = 171$ $\quad x_{max} = 408$

simple box plot



modified box plot

$f_s = f_3 - f_1 = 252.5 - 202 = 50.5$

$IF = [f_1 - 1.5f_s \quad f_3 + 1.5f_s] = [202 - 1.5 \times 50.5 \quad 252.5 + 1.5 \times 50.5]$

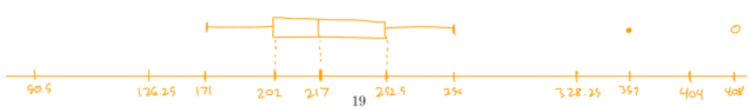
$= [126.25, \ 328.25]$

$OF = [f_1 - 3f_s \quad f_3 + 3f_s] = [202 - 3 \times 50.5 \quad 252.5 + 3 \times 50.5]$

$= [50.5 \quad 404]$

mild outlier: 359

extreme outlier: 408



Chapter 2 Stuff

Conditional probability of A given that the event B has occurred:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0.$$

Multiplication rule:

$$P(A \cap B) = P(A|B) * P(B).$$

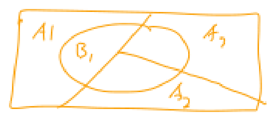
The law of total probability: Let $A_1, A_2, ..., A_k$ be mutually exclusive and exhaustive events, then for any other event B.

$$P(B) = \sum_{i=1}^{k} P(B|A_i)*P(A_i) = P(B|A_1)*P(A_1)+P(B|A_2)*P(A_2)+...+P(B|A_k)*P(A_k).$$

Because $A_1, A_2, ..., A_k$ are disjoint and exhaustive

$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_k \cap B)$

$P(B) = \sum_{i=1}^{k} P(A_i \cap B) \xRightarrow[\text{rule}]{\text{multiplication}} \sum_{i=1}^{k} P(B/A_i) \cdot P(A_i)$



Baye's Theorem Let $A_1, A_2, ..., A_k$ be a collection of k mutually exclusive and exhaustive events, with prior probability $P(A_i), i = 1, 2, ..., k$. Then for any other event B for which $P(B) > 0$, the posterior probability of A_j given that B has occurred is

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} \xrightarrow{\text{multiplication rule}} = \frac{P(B|A_j) * P(A_j)}{\sum_{i=1}^{k} P(B|A_i) * P(A_i)}.$$

→ the law of total probability

Chapter 3 Stuff

3.3 Expected value of X

3.3.1 Expected value

Example: Bernoulli $D = \{0,1\}$ $P(0)=0.5$ $P(1)=0.5$ $E(x) = 0\times0.5 + 1\times0.5 = 0.5$

Let X be a discrete rv with set of possible values D and pmf $p(x)$. the **expected value** or **mean value** of X, denoted by $E(X)$ or μ_X:

$$E(X) = \mu_X = \sum_{x \in D} \overset{\text{value}}{x} * \overset{P(x=x)}{p(x)}.$$

The **expected value** of any function $h(X)$, denoted by $E[h(X)]$ or $\mu_{h(X)}$:

$$E[h(X)] = \sum_{x \in D} h(x) * p(x).$$

Question: what is the different between $\bar{x}$, μ and μ_X?

$\bar{x}$: sample mean; average of a given sample

μ: population mean; average of the entire population

μ_x: mean value of

3.3.2 Variance of X

σ^2: variance of data

Let X be a discrete rv with pmf $p(x)$, and expected value μ. Then the **variance** of X, denoted by $V(X)$ or σ_X^2, or σ^2 is

$$V(X) = \sum_{x \in D} (x - \mu)^2 * p(x) = E[(X - \mu)^2], \text{ or}$$

$$V(X) = \sum_{x \in D} x^2 * p(x) - \mu^2 = E(X^2) - [E(X)]^2. \text{ → very useful}$$

The **standard deviation (SD)** of X is $\sigma_X = \sqrt{\sigma_X^2}$.

The variance of function $h(X)$:

$$V[h(X)] = \sigma_{h(X)}^2 = \sum_{x \in D} (h(x) - E[h(X)])^2 * p(x) = E[(h(X) - E(h(X)))^2].$$

Binomial random variable, X, is defined as the number of success in n trials.

And the probability of success is denoted by p, the pmf of X:

$$X \sim b(x, n, p).$$

$$b(x,n,p) = p(x) = p(X=x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!\,(n-x)!} \cdot p^x (1-p)^{n-x}$$

$x = 0, 1, 2 \ldots n-1, n$

Mean of X: expected value of x: $n \cdot p$

Variance of X: $n \cdot p \cdot (1-p)$

Standard deviation of X: $\sqrt{n \cdot p(1-p)}$

Cumulative probability for a binomial random variable, X, is defined as:

3.5 The Poisson probability distribution

The **poisson distribution** is often used to count rare events.
discrete distribution

$P(x \le 4) = P(0) + P(1) + \underbrace{P(2) + P(3) + P(4)}_{P(2 \le x \le 4)}$

$P(x \le 1) = P(0) + P(1)$

Poisson experiment:

1. The probability of a single event occurs in a given interval (of time, length, volumn...) is the same for all interval. = 3 / hour

2. The number of events that occur in any interval is independent of others.

Poisson random variable, X, is a count of the number of times the specific event occurs during a given interval.

e: euler # $e \approx 2.71828$

$x: 0, 1, 2, 3, 4 \ldots$

The pmf: $P(X = x) = \dfrac{e^{-\mu} \cdot \mu^{x}}{x!}$

The mean: $E(x) = \mu$

The variance: $V(x) = \mu$

The cdf: Appendix A2

The Poisson distribution as a limit of Binomial

Suppose that in the binomial pmf $b(x; n, p)$, we let $n \longrightarrow \infty$ and $p \longrightarrow 0$. In such a way that np approaches a value $\mu > 0$, then

$$b(x; n, p) \xrightarrow{\text{approximate}} p(x; \mu)$$

This approximation can safely be applied if $n > 50$ $\quad n \cdot p < 5$ → p to be small

Poisson process

Let α be the average event occurring rate in a unit time period.

Let rv X be the number of events occurs during any time interval of length t.

$X \sim p(x; \alpha t)$.

$\alpha t = 3 \times 24 = 72 / \text{day}$

$x \sim p(x, 72)$

Chapter 4 Stuff:

4.1 Probability density functions

A **continuous probability distribution** completely describes the random variable and is used to compute probabilities associated with random variable.

Probability density function (pdf), $f(x)$:

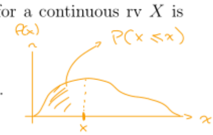
1. is a function defined for all real numbers. i.e. $x \in (-\infty, +\infty)$.

2. is a smooth curve describes the **probability distribution** for a continuous random variable X through area under the curve. Let $a \le b$, the probability

$$P(a \le X \le b) = \int_a^b f(x)dx.$$ Is $f(a) = p(x=a)$?

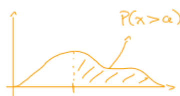
The **cumulative distribution function (cdf)** $F(x)$ for a continuous rv X is defined for every number x by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy.$$



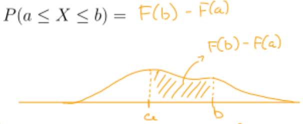
$F(x)$ is the area under the density curve to the left of x.

Note: For any numbers $a, b, a \le b$.



1. $P(X > a) =$ $1 - P(x \le a) = 1 - F(a)$

2. $P(a \le X \le b) =$ $F(b) - F(a)$

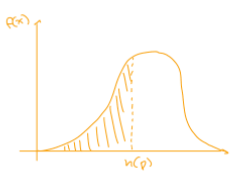


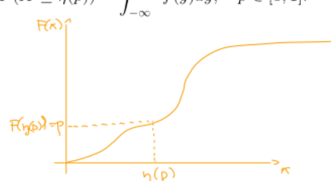
pdf

3. $f(x) = F'(x)$. $\Longleftrightarrow$ $F(x) = \int_{-\infty}^{x} f(x)dx$

4. The $(100p)$th percentile, $\eta(p)$, is defined by

$$p = F(\eta(p)) = P(X \le \eta(p)) = \int_{-\infty}^{\eta(p)} f(y)dy, \quad p \in [0, 1].$$

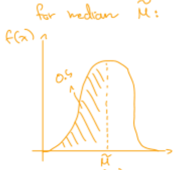


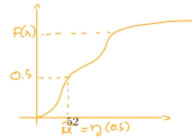


5. The **median** of a continuous distribution, denoted by $\tilde{\mu}$, is the 50th percentile.

so $F(\tilde{\mu}) = 0.5$. Half the area under the pdf curve is to the left of μ.

for median $\tilde{\mu}$: $p = 0.5$ $\tilde{\mu} = \eta(0.5)$ $F(\tilde{\mu}) = F(\eta(0.5)) = 0.5$





6. The **Expected value** and variance of a continuous rv X:

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

$\mu_{h(X)} = E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x)\, dx$

$\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x - E(x))^2 f(x)\, dx$

$\sigma_X = \sqrt{\sigma_x^2}$

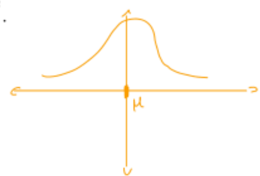
4.2 The normal distribution

Family of distributions

Normal distribution has two parameters: μ, σ (or μ, σ^2), $-\infty < \mu < +\infty$ and $\sigma > 0$. We write the random variable $X \sim N(\mu, \sigma^2)$. The pdf is:

$p(x<1) = \int_{-\infty} f(x)\, dx \rightarrow$ uneasy

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$



1. Bell shaped, unimodal

2. Symmetric about $x = \mu$

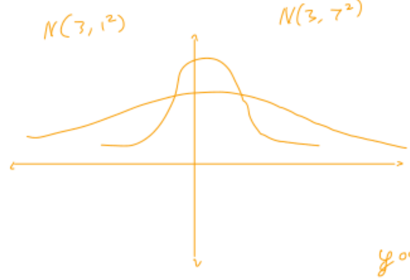
3. σ (σ^2) represent the spread $v(x) = \sigma^2$

Side question: How many normal distributions are there?
Infinitely many

Example:

$N(7, 5^2)$ $N(3, 5^2)$

$N(3, 1^2)$ $N(3, 7^2)$



you may have missed something
Mar 5

specific normal distribution

rv: $Z \to N(0,1)$

Standard normal distribution $N(0,1)$: That is $\mu = 0, \sigma = 1$. The pdf is

$$f(z;0,1) = \frac{1}{2\pi}e^{-z^2/2}.$$

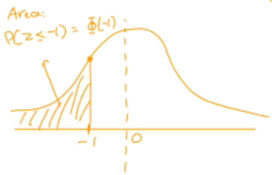
cdf: $F(x)$

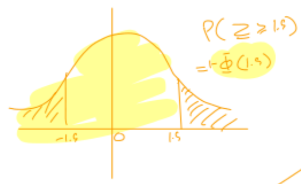
The cdf: Appendix A.3

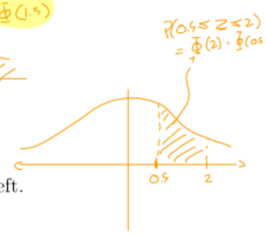
value

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} f(y)dy = \int_{-\infty}^{z} \frac{1}{2\pi}e^{-z^2/2}dy$$

$P(X \leq x)$



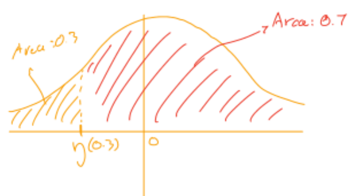


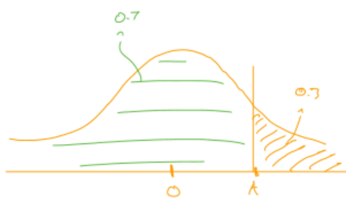


(100p)th percentile vs critical value z_α

- **(100p)th percentile** $\eta(p)$: p refers to the area on the left.
- **Critical value** z_α: α refers to the area on the right.

Example: $P = 0.3$ $\eta(0.3)$





$\underline{A}$: $\eta(0.7)$

$Z_{0.3}$ $\alpha = 0.3$

$\eta(0.3) = \underline{Z_{0.7}}$

30th percentile

Z-critical value with $\alpha = 0.7$

Standardization a normal rv X: $X \sim N(\mu, \sigma^2)$

$$X \sim N(\mu, \sigma^2) \xrightarrow{\text{Standardization}} Z \sim N(0,1)$$

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

$$P(x \leq a) = P\Big(\underbrace{\frac{x-\mu}{\sigma}}_{\text{new rv. } \sim N(0,1)} \leq \frac{a-\mu}{\sigma}\Big) = P\Big(Z \leq \frac{a-\mu}{\sigma}\Big)$$

$$= \Phi\Big(\frac{a-\mu}{\sigma}\Big)$$

* $x \leq a$ and $\frac{x-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}$ are equivalent

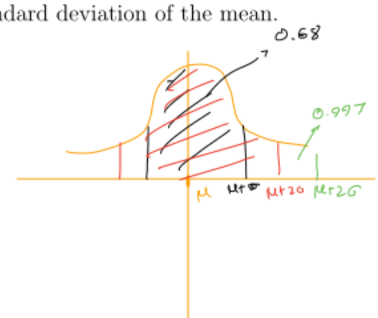
The empirical rule: If the population distribution of a variable is (approximately) normal, then

1. Roughly 68% of the values are within 1 standard deviation of the mean.
2. Roughly 95% of the values are within 2 standard deviation of the mean.
3. Roughly 99.7% of the values are within 3 standard deviation of the mean.

$P(\mu-\sigma \leq x \leq \mu+\sigma) \approx 0.68$

$P(\mu-2\sigma \leq x \leq \mu+2\sigma) \approx 0.95$

$P(\mu-3\sigma \leq x \leq \mu+3\sigma) \approx 0.997$



Approximating the binomial distribution

$= 0.1717$

Let X be a binomial rv based on n trials with success probability p, so $X \sim b(n, p)$.

Let Y be a normal rv, $\mu = n * p$, $\sigma = \sqrt{np(1-p)}$, $Y \sim N(\mu, \sigma^2)$.

When $np \geq 10$ and $n(1-p) \geq 10$,

$$P(X \leq x) = b(x; n, p) \approx P(Y \leq x + 0.5).$$

Chapter 5 Stuff

5.2 The distribution of the sample mean

Proposition 1 The rv's $X_1, X_2, ..., X_n$ be a random sample from a distribution with mean value μ and standard deviation σ. Let $T_0 = X_1 + X_2 + ... + X_n$, the sample total. Then

1. $E(\bar{X}) = \mu$. The mean of sample mean is the same as population mean
2. $V(\bar{X}) = \sigma^2/n$, $\sigma_{\bar{X}=\sigma/\sqrt{n}}$. The variance of sample mean is smaller than population variance
3. $E(T_0) = n\mu$, $V(T_0) = n\sigma^2$, $\sigma_{T_0} = \sqrt{n}\sigma$.

Proposition 2 The rv's $X_1, X_2, ..., X_n$ be a random sample from a **normal** distribution with mean value μ and standard deviation σ, then for any n

1. $E(\bar{X}) = \mu$, $V(\bar{X}) = \sigma^2/n$. same as above
2. $\bar{X} \sim N(\mu, \sigma^2/n)$. normal distribution for $\bar{x}$
3. $T_0 \sim N(n\mu, n\sigma^2)$.

Central limit theorem (CLT)

The rv's $X_1, X_2, ..., X_n$ be a random sample from a distribution with mean value μ and standard deviation σ. If n is sufficiently large $(n > 30)$,then

1. $E(\bar{X}) = \mu$, $V(\bar{X}) = \sigma^2/n$.
2. $\bar{X} \sim N(\mu, \sigma^2/n)$
3. $T_0 \sim N(n\mu, n\sigma^2)$

Example (exercise 47 P237) $X \sim N(70, 1.6^2)$, $n = 16$, find $P(69 \le \bar{X} \le 71)$.

$$\left.\begin{array}{l} X \sim N(70, 1.6^2) \\ n=16 \end{array}\right\} \Rightarrow \bar{X} \sim N\left(70, \frac{1.6^2}{16}\right)$$

$$P(69 \le \bar{X} \le 71) = P\left(\frac{69-7}{1.6/4} \le Z \le \frac{71-70}{\frac{1.6}{4}}\right)$$

$$= P(-2.5 \le Z \le 2.5) = 0.9938 - 0.0062 = 0.9876$$

$$\sigma^2 = \frac{1.6^2}{16} \qquad \sigma = \sqrt{\frac{1.6^2}{16}} = \frac{1.6}{4}$$

Chapter 6 Stuff

To estimate the population mean μ, we an choose the following point estimators.

μ: population parameter, denote as Θ,
want to find point estimator $\hat{\Theta}$ in this case is $\hat{\mu}$ —> sample measurement

① $\hat{\mu} = \bar{x}$ sample mean —> point estimation

$\bar{x} = \frac{1}{20} \sum_{i=1}^{20} x_i = 27.793$ —> point estimate

② $\hat{\mu} = \tilde{x}$ sample median —> point estimator

$\tilde{x} = \frac{27.94 + 27.98}{2} = 27.96$ —> point estimate

③ $\hat{\mu} = \frac{x_{max} - x_{min}}{2}$ —> point estimator $\frac{24.46 + 30.88}{2} = 27.67$ } Point estimate

P: population proportion

point estimator: $\hat{p} = \frac{x}{n}$ —> sample proportion

point estimate: $\hat{p} = \frac{x}{n} = \frac{15}{25} = 0.6$ —> calculated by the sample

A point estimator $\hat{\theta}$ is said to be

Which is better? unbiased provides a better estimate

1. **unbiased estimator** of θ, if $E(\hat{\Theta}) = \Theta$
1. **biased estimator** of θ, if $E(\hat{\Theta}) \neq \Theta$

Some unbiased estimator:

parameter	unbiased estimator	Comments
$p : X \sim b(n, p)$	$\hat{p} = \frac{x}{n}$	$E(\frac{x}{n}) = \frac{1}{n} E(x) = \frac{1}{n}(n \cdot p) = p$
$\mu = \sum_{i=1}^{N} X_i/N$	$\hat{\mu} = \bar{x}$	$E(\bar{x}) = E(\frac{\sum x_i}{n}) = \frac{1}{n} E(\sum x_i) = \frac{1}{n} \sum E x_i = \frac{1}{n} \sum \mu$
$\mu = \sum_{i=1}^{N} X_i/N$	$\hat{\mu} = \tilde{x}$	if the distribution is continuous and symmetric
$\mu = \sum_{i=1}^{N} X_i/N$	$\hat{\mu} = x_{tr(p)}$	
$\sigma^2 = \sum_{i=1}^{N}(X_i - \mu)^2/N$	$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$	$E(S^2) = \sigma^2$

Minimum variance unbiased estimation (MVUE): Among all estimators that are unbiased, the one that has minimum variance is called **MVUE**.

Example

$\bar{X}$ is MVUE of μ

$$\left.\begin{array}{l} V(\bar{x}) \leq V(\tilde{x}) \\ V(\bar{x}) \leq V(\bar{x}_{tr(p)}) \end{array}\right\}$$

is MVUE good?

MVUE is better smaller variance for an estimation is good

Standard error of an estimator $\hat{\theta}$: describes the magnitude of a typical or representative deviation between an estimate and the true value of θ

$\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})} = \sqrt{\sigma^2_{\hat{\theta}}} \longrightarrow$ Standard deviation of $\hat{\theta}$

Estimated standard error of an estimator $\hat{\theta}$: If the standard error of the estimator itself involves unkown parameters, whose value can be estimated.

denoted by $\hat{\sigma}_{\theta}$ or $S_{\hat{\theta}}$ means the standard error involves variables

Example 6.9 P259 A normal distribution $N(\mu, \sigma)$.

S^2 = sample variance

$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$

$\bar{x}$ is estimator of μ (MVUE) let $n=20$

$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ by proposition 2

$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} \longrightarrow$ Standard error of estimation

① if σ is Known $\sigma_{\bar{x}} = \frac{1.5}{\sqrt{20}} = 0.335 \rightarrow$ standard error

② if σ is unknown $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{20}}$ involves a variable

use sample sd s to estimate σ, let's say $s = 1.462$

$\hat{\sigma}_{\bar{x}} = \frac{S}{\sqrt{20}} = \frac{1.462}{\sqrt{20}} = 0.32 \rightarrow$ estimated standard error

6.2.1 The method of moments

Let $X_1, X_2, ..., X_n$ be a random sample from a pmf or pdf $f(x)$. For $k = 1, 2, 3, ...$

the k**th population moment**(k**th moment of the distribution** $f(x)$): $E(X^k)$.

the k**th sample moment**: $\frac{1}{n}\sum_{i=1}^{n} X_i^k$.

	1st	2nd	3rd	
Population moment	$E(x)$	$E(x^2)$	$E(x^3)$	
Sample moment	$\frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$	$\frac{1}{n}\sum_{i=1}^{n} x_i^2$	$\frac{1}{n}\sum_{i=1}^{n} x_i^3$	....

Example:

Let $X_1, X_2, ..., X_n$ be a random sample from a pmf or pdf $f(x; \theta_1, ..., \theta_m)$, where $\theta_1, ..., \theta_m$ are parameters whose values are unknown.

the **Moment estimators** $\hat{\theta}_1, ..., \hat{\theta}_m$ are obtained by equating the first m sample moments to the corresponding first m population moments and solving for $\theta_1, ..., \theta_m$.

Example 6.12 P265 Let $X_1, X_2, ..., X_n$ be a random sample from exponential distribution. Find the moment estimator of parameter λ. $X \sim Exp(\lambda)$

1st P.M $E(x) = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{E(x)}$

1st S.M $\bar{x}$

hence $\hat{\lambda} = \frac{1}{\bar{x}}$ by equating $E(x)$, $\bar{x}$

6.2.2 Maximum likelihood estimation

Let $X_1, X_2, ..., X_n$ be a random sample from a pmf or pdf $f(x; \theta_1, ..., \theta_m)$, $x_1, x_2, ..., x_n$ are the observed sample values. $x_1, x_2, x_3 \cdots x_m$ are variables

The **joint pmf/pdf**: $\theta_1, \theta_2, \theta_3, \ldots \theta_m$ are given

$$f(x_1, x_2, ..., x_n; \theta_1, ..., \theta_m) = \overbrace{\prod_{i=1}^{n}}^{\text{Product}} f(x_i; \theta_1, ..., \theta_m)$$

$$= f(x_1; \underbrace{\theta_1, ..., \theta_m}_{\text{variables}}) f(\underbrace{x_2}_{\text{given}}; \theta_1, ..., \theta_m) ... f(x_n; \theta_1, ..., \theta_m)$$

The **likelihood function**:

$$f(x_1, x_2, ..., x_n; \theta_1, ..., \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, ..., \theta_m)$$

$$= f(x_1; \theta_1, ..., \theta_m) f(x_2; \theta_1, ..., \theta_m) ... f(x_n; \theta_1, ..., \theta_m)$$

The **natural logarithm of likelihood function**:

$$\ln f(x_1, x_2, ..., x_n; \theta_1, ..., \theta_m) = \sum_{i=1}^{n} \ln f(x_i; \theta_1, ..., \theta_m)$$

$$= \ln f(x_1; \theta_1, ..., \theta_m) + \ln f(x_2; \theta_1, ..., \theta_m) + ... + \ln f(x_n; \theta_1, ..., \theta_m)$$

The **maximum likelihood estimators (mle)** $\hat{\theta}_1, ..., \hat{\theta}_m$: Those values of θ_i's that maximize the likelihood function.

Choose MLE $\hat{\theta}_1, \hat{\theta}_2 \ldots \hat{\theta}_m$ to make observed $x_1, x_2, \ldots x_m$ most likely

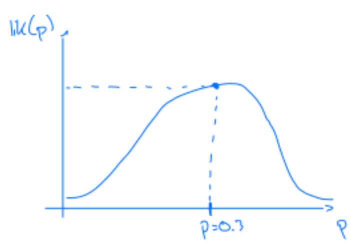
Hence $P(X = 1) = p, P(X = 0) = 1 - p$. Suppose 10 individuals are randomly selected. $x_1 = x_3 = x_{10} = 1$, other seven x_i's are all zero. Find the mle of p.

$Lik(p) = P(x_1 = 1) \cdot P(x_2 = 0) \cdot P(x_3 = 1) \cdot \ldots \cdot P(x_{10} = 1) = p(1-p) \cdot p \cdots p$

$$lik(p) = \begin{cases} p^3(1-p)^7 = (0.25^3)(0.75^7) = 0.002086 \\ p^3(1-p)^7 = (0.5^3)(0.5^7) = 0.000977 \end{cases}$$

$p = 0.25$ make the sample more likely

$p = 0.5$ $\quad p \in [0,1]$



$\ln(lik(p)) = \ln[p^3(1-p)^7]$

$= \ln(p^3) + \ln(1-p)^7 = 3\ln p + 7\ln(1-p)$

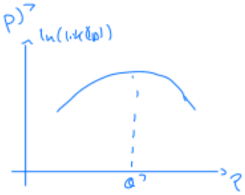
$$\frac{\partial \ln(lik(p))}{\partial p} = \frac{3}{p} + \frac{7}{1-p} \times (-1)$$

$$= \frac{3}{p} - \frac{7}{1-p} = 0$$

$p = 0.3$ maximizes $\ln(lik(p))$

also maximizes $lik(p)$

when $p = 0.3$ $lik(p) = p^3(1-p)^7$ has the largest value



$\hat{p} = 0.3$ is mle

Chapter 7 Stuff

7.1.4 The width of CI and sample size

Given the formula of $100(1-\alpha)\%$ confidence interval $\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. The width is

$$W = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

We can then solve for n, which will derive the desired CI width.

$$n = \left(2z_{\alpha/2}\frac{\sigma}{W}\right)^2$$

Note: The following factors may effect the width of a CI.

1. $\alpha \downarrow \rightarrow CL \uparrow \rightarrow z_{\frac{\alpha}{2}} \uparrow \rightarrow CI \uparrow$

2. $\sigma \uparrow \rightarrow CI \uparrow$

3. $n \uparrow \rightarrow CI \downarrow$

7.2.4 One-sided CI (confidence bounds)

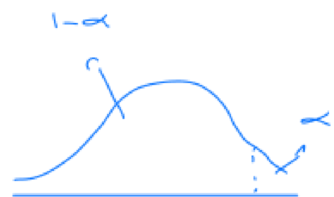
Large sample one-sided CI gives upper bound for μ:

→ Large Sample $n>40$

→ σ unknown

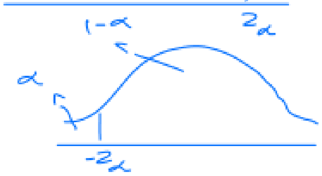
→ Population's distribution is unknown

$$(-\infty, \bar{x} + z_\alpha\frac{s}{\sqrt{n}}).$$



Large sample one-sided CI gives lower bound for μ:

$$(\bar{x} - z_\alpha\frac{s}{\sqrt{n}}, \infty).$$



7.3.3 Two-sided CI under different assumptions

	population	sample size	variance σ	distribution	CI
1	Normal	-	known	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$[\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$
2	-	Large ($n \geq 40$)	known	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$[\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$
3	-	Large ($n \geq 40$)	unknown	$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$	$[\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}]$
4	Normal	Small ($n \leq 40$)	unknown	$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$	$[\bar{x} \pm t^{n-1}_{\alpha/2}\frac{s}{\sqrt{n}}]$

7.3.4 A prediction interval for a single future value

A random sample $X_1, X_2, ..., X_n$ from a normal distribution, we would like to predict the value to be observed at some future time, X_{n+1}.

A point predictor of X_{n+1} is $\bar{X}$. Hence the predict error is: $\bar{x} - x_{n+1}$

$\bar{x}$ prediction, x_{n+1} true value

1. The expected value of predict error:

$E(\bar{x} - x_{n+1}) = E(\bar{x}) - E(x_{n+1}) = \mu - \mu = 0$

2. The variance of predict error:

$V(\bar{x} - x_{n+1}) = V(\bar{x}) + V(x_{n+1}) = \frac{\sigma^2}{n} + \sigma^2 = \sigma^2(1 + \frac{1}{n})$

$\bar{x}, x_{n+1}$ are independent $V(x \pm y) = V(x) + V(y)$ when x, y are independent

3. The distribution of the predict error:

$\bar{x}, x_{n+1}$ are both normal, and independent

$\bar{x} - x_{n+1}$ is also normal

$\bar{x} - x_{n+1} \sim N(0, \sigma^2(1 + \frac{1}{n}))$

$$\frac{(\bar{x} - x_{n+1}) - 0}{\sigma\sqrt{1 + \frac{1}{n}}} \sim N(0,1)$$

A prediction interval (PI) for a single observation to be selected from a normal population is

if σ is unknown, we will use s to replace σ

$$= \frac{(\bar{x} - x_{n+1}) - 0}{\frac{s}{\sqrt{n}}} \sim t(n-1)$$

2 sided: the $100(1-\alpha)\%$ P.I $\bar{x} \pm t_{\frac{\alpha}{2}}(n-1)\, s \cdot \sqrt{1 + \frac{1}{n}}$

1 sided: P.I with upper bound $(-\infty, \bar{x} + t_{\alpha}(n-1)(s)\sqrt{1 + \frac{1}{n}})$

lower bound $(\bar{x} - t_{\alpha}(n-1)\, s\sqrt{1 + \frac{1}{n}}, +\infty)$

Chapter 8 Stuff

We now show how to solve it using P-**value**:

1. Write down the hypotheses:

$H_0: \mu = 76$ $\quad$ $H_a: \mu > 76$

2. Figure out the distribution of $\bar{X}$, calculate the P-value $P(\bar{X} \geq \bar{x} | H_0 \text{ is true})$

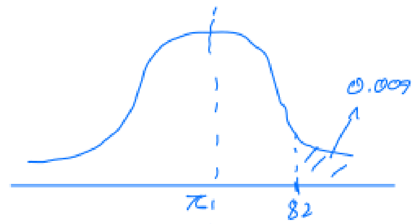
X: weight of canadian males (30-44) $X \sim N(\mu, 25^2)$

$\bar{x}$: sample mean $\bar{X} \sim N(\mu, \frac{25^2}{n}) = N(\mu, (\frac{25}{10})^2)$

Assume H_0 is true, calculate P-Value

$P(\bar{X} \geq \bar{x} | H_0 \text{ is true}) = P(\bar{X} \geq 82 | \mu = 76)$

$= P(\frac{\bar{x} - 76}{\frac{25}{10}} > \frac{82-76}{\frac{25}{10}}) = P(Z \geq 2.4) = 0.009$

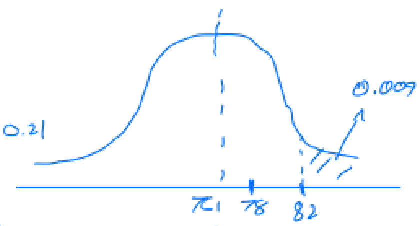


If H_0 is true, it is very unlikely ($p = 0.009$) to get a sample mean of 82 or above. But now, a rare event ($\bar{x} = 82$) occurs hence there is evidence that H_0 is wrong

3. Choose a threshold α, compare with P-value.

What if $\bar{x} = 78$, calculate p-value

$P(\bar{X} \geq 78 | H_0 \text{ is true}) = P(\frac{\bar{x}-76}{\frac{25}{10}} \geq \frac{78-76}{\frac{25}{10}}) = P(Z \geq 0.8) = 0.21$



Is $\bar{x} = 78$ a rare event or a reasonable event?

Hence, we will set a threshold α (significance level)

4. Make conclusion.

Choose small values for α Example: $\alpha = 0.01, 0.05, 0.1$

Compare p-value with α

P-value $\leq \alpha$ rare event (unlikely to occur), reject H_0 in favor of H_a

P-value $> \alpha$ reasonable event (likely to occur) accept H_0

We then show how to solve it using **reject point, reject region (RR)**:

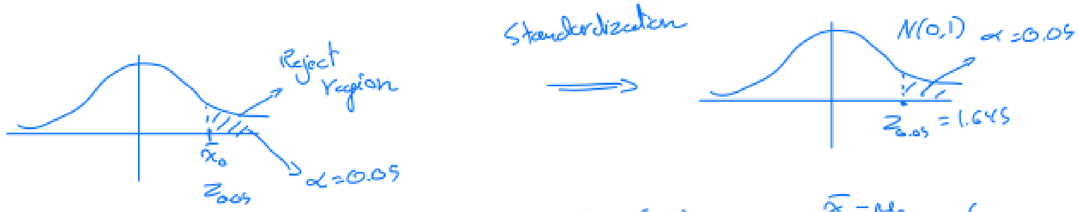
1. Write down the hypotheses:

$H_0: \mu = 76 \qquad H_a: \mu > 76$

2. Figure out the distribution of $\bar{X}$, determine/calculate the test statistic (TS).

$X \sim N(\mu, \sigma^2) \implies \bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

Start to assume H_0 is true, $\mu_0 = \mu = 76$ $\quad \bar{x} \sim N(76, (\frac{25}{10})^2)$



Define test statistic (TS): $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ (z-score of $\bar{x}$)

3. Choose a threshold α, calculate reject point and reject region (RR).

ussually α small value 0.05, 0.01, 0.1

In this case we choose $\alpha = 0.05$, we want to calculate the cut off (reject point) $\bar{x}_0$

such that if H_0 is true, then 95% sample mean will be less than $\bar{x}_0$

That is $\frac{\bar{x}_0 - \mu_0}{\frac{\sigma}{\sqrt{n}}} = Z_\alpha = Z_{0.05} = 1.645$, then we can solve $\bar{x}_0$, then compare

$\bar{x} = 82$ with $\bar{x}_0$

Note: compare $\bar{x}$ with $\bar{x}_0$ is equivalent to compare TS $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ with Z_α

Define: Reject Point as Z_α

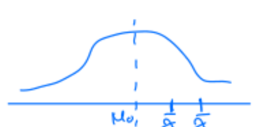
Reject Region as $(Z_\alpha, +\infty)$

4. Make conclusion

TS
① $z \geqslant Z_\alpha$ (ie: $z \in (Z_\alpha, +\infty)$): rare event → reject H_0
② $z < Z_\alpha$ (ie: $z \notin (Z_\alpha, +\infty)$) reasonable → accept H_0

①

②

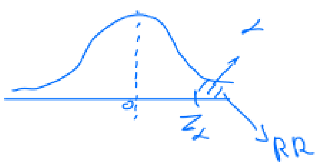


8.1.2 Three types alternative hypothesis

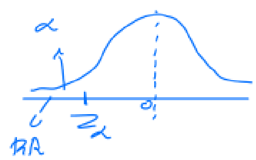
Upper one-tailed, lower one-tailed, two-tailed

$H_0: \mu = \mu_0$

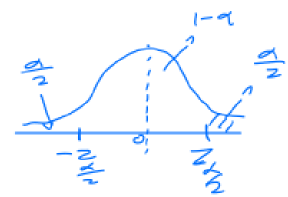
$H_a: \mu > \mu_0$ | $H_a: \mu < \mu_0$ | $H_a: \mu \neq \mu_0$ ($\mu < \mu_0$, $\mu > \mu_0$)



Upper



lower



RR: $(-\infty, -Z_{\frac{\alpha}{2}}) \cup (Z_{\frac{\alpha}{2}}, +\infty)$

Check both tails

8.1.5 Hypothesis test - normal distribution, unknown σ, and small sample

Recall: when

1. Underlying distribution is normal, and
2. Unknown σ^2. We have

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

Hence, When the underlying population is normal, but the population variance σ^2 is unknown. To implement the hypothesis test, we define:

Test statistic (TS): $T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ $\quad t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

z_α $\quad -z_\alpha$ $\quad \pm z_{\frac{\alpha}{2}}$

	$H_a: \mu > \mu_0$	$H_a: \mu < \mu_0$	$H_a: \mu \neq \mu_0$
Reject Point	$t_\alpha(n-1)$	$-t_\alpha(n-1)$	$\pm t_{\frac{\alpha}{2}}(n-1)$
RR	$(t_\alpha(n-1), +\infty)$	$(-\infty, -t_\alpha(n-1))$	$(-\infty, -t_{\frac{\alpha}{2}}(n-1)) \cup (t_{\frac{\alpha}{2}}(n-1), +\infty)$
P-value	$P(T \geq t)$	$P(T \leq t)$	$2P(T \geq t)$

We reject H_0, when:

1. P-Value $\leq \alpha$

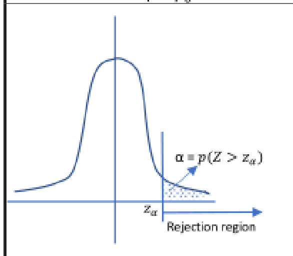
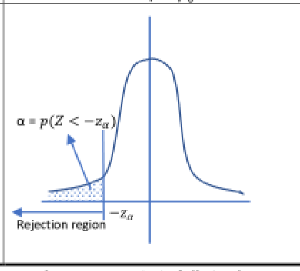
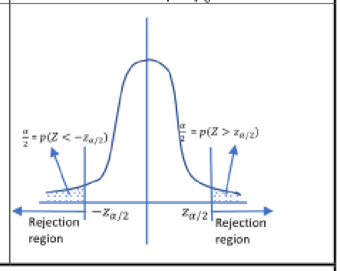
or

2. TS: $t \in RR$

Determine test statistic for different assumptions

	population	sample size	variance σ	statistic	distribution	
1	Normal	-	known	$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	
2	Normal	Small ($n \leq 40$)	unknown	$t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$	$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$	
3	-	Large ($n \geq 40$)	known	$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	CLT
4	-	Large ($n \geq 40$)	unknown	$z = \frac{\bar{x}-\mu}{s/\sqrt{n}}$	$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$	

Summary of solving hypothesis tests:

<table>
<tr><td colspan="3">H0 : $\mu = \mu_0$</td></tr>
<tr><td>Upper-tailed test</td><td>Lower-tailed test</td><td>Two-tailed test</td></tr>
<tr><td>Ha: $\mu > \mu_0$</td><td>Ha: $\mu < \mu_0$</td><td>Ha: $\mu \neq \mu_0$</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td colspan="3">Solve using Reject region (RR): reject H0, when test statistic falls in the RR.</td></tr>
<tr><td colspan="3">Test statistic: $Z \sim N(0,1)$</td></tr>
<tr><td>Reject point: z_α</td><td>Reject point: $-z_\alpha$</td><td>Reject point: $\pm z_{\alpha/2}$</td></tr>
<tr><td>RR: (z_α, ∞)</td><td>RR: $(-\infty, -z_\alpha)$</td><td>RR: $(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$</td></tr>
<tr><td colspan="3">Test statistic: $T \sim t(n-1)$</td></tr>
<tr><td>Reject point: $t_\alpha(n-1)$</td><td>Reject point: $-t_\alpha(n-1)$</td><td>Reject point: $\pm t_{\alpha/2}(n-1)$</td></tr>
<tr><td>RR: $(t_\alpha(n-1), \infty)$</td><td>RR: $(-\infty, -t_\alpha(n-1))$</td><td>RR: $(-\infty, -t_{\alpha/2}(n-1)) \cup (t_{\alpha/2}(n-1), \infty)$</td></tr>
<tr><td colspan="3">Solve using P-value: reject H0, when P-value $\leq \alpha$.</td></tr>
<tr><td colspan="3">Test statistic: $Z \sim N(0,1)$</td></tr>
<tr><td>P-value = $P(\bar{X} \geq \bar{x} \mid$ H0 is true)
= $P(Z \geq z)$</td><td>P-value= $P(\bar{X} \leq \bar{x} \mid$ H0 is true)
= $P(Z \leq z)$</td><td>P-value=2 $* P(\bar{X} \geq \lvert\bar{x}\rvert \mid$ H0 is true)
= 2*$P(Z \geq \lvert z\rvert)$</td></tr>
<tr><td colspan="3">Test statistic: $T \sim t(n-1)$</td></tr>
<tr><td>P-value = $P(\bar{X} \geq \bar{x} \mid$ H0 is true)
= $P(T \geq t)$</td><td>P-value= $P(\bar{X} \leq \bar{x} \mid$ H0 is true)
= $P(T \leq t)$</td><td>P-value=2 $* P(\bar{X} \geq \lvert\bar{x}\rvert \mid$ H0 is true)
= 2*$P(T \geq \lvert t\rvert)$</td></tr>
<tr><td colspan="3">Solve using confidence interval: reject H0, when $\mu_0 \notin$ CI</td></tr>
<tr><td colspan="3">Test statistic: $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$</td></tr>
<tr><td>$(\bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}, \ +\infty)$</td><td>$(-\infty, \ \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}})$</td><td>$(-z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}})$</td></tr>
<tr><td colspan="3">Test statistic: $Z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim N(0,1)$</td></tr>
<tr><td>$(\bar{x} - z_\alpha \frac{S}{\sqrt{n}}, \ +\infty)$</td><td>$(-\infty, \ \bar{x} + z_\alpha \frac{S}{\sqrt{n}})$</td><td>$(\bar{x} - z_{\alpha/2} \frac{S}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2} \frac{S}{\sqrt{n}})$</td></tr>
<tr><td colspan="3">Test statistic: $T \sim t(n-1)$</td></tr>
<tr><td>$(\bar{x} - t_\alpha^{(n-1)} \frac{S}{\sqrt{n}}, \ +\infty)$</td><td>$(-\infty, \ \bar{x} + t_\alpha^{(n-1)} \frac{S}{\sqrt{n}})$</td><td>$(\bar{x} - t_{\alpha/2}^{(n-1)} \frac{S}{\sqrt{n}}, \ \bar{x} + t_{\alpha/2}^{(n-1)} \frac{S}{\sqrt{n}})$</td></tr>
</table>

Example 8.5 P319 The drying time of a type of paint under specified test conditions is known to be normally distributed with mean value 75 mins and standard deviation 9 mins. A new additive is designed to decrease average drying time. It is believed that drying time with this additive will remain normally distributed with $\sigma = 9$. We want to test if this additive will decrease the drying time. $\alpha = 0.01$, $n = 25$. Compute the probability of type II error $\beta(\mu_a)$.

$H_0: \mu = 75 \qquad H_a: \mu < 75$

normal, σ unknown $\Big] \Rightarrow$ TS: $Z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x} - 75}{\frac{9}{\sqrt{25}}} = \frac{\bar{x} - 75}{1.8}$$

To reject H_0 we need p-value $< \alpha$, or R.P to be $-Z_\alpha = -Z_{0.01} = -2.325$

That is equivalent to $\dfrac{\bar{x}_0 - \mu_0}{\frac{\sigma}{\sqrt{n}}} = -2.325 \Rightarrow \bar{x}_0 = 70.8$

Type II error: Accept H_0 given H_a is true

assume the true mean $\mu_a = 72 < 75$

① $\beta(72) = P(\text{Accept } H_0 \mid \mu_a = 72)$
$= P(\bar{x} > 70.8 \mid \bar{x} \sim N(72, 1.8^2))$
$= P\left(Z > \dfrac{70.8 - 72}{1.8}\right) = P(Z > -0.67) = 0.7486$

μ_a is further from μ_0 then $\beta(\mu_a) \downarrow$

② $\beta(70) = P(\bar{x} > 70.8 \mid \bar{x} \sim N(70, 1.8^2))$
$= P\left(Z > \dfrac{70.8 - 70}{1.8}\right) = P(Z > 0.44) = 0.33$

③ $\beta(67) = P(\bar{x} > 70.8 \mid \bar{x} \sim N(67, 1.8^2))$
$= P(Z > 2.11) = 0.174$

Alternative Hypothesis	Type II Error Probability $\beta(\mu')$ for a Level α Test
H_a: $\mu > \mu_0$	$\Phi\left(z_\alpha + \dfrac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \dfrac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
H_a: $\mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \dfrac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \dfrac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

where $\Phi(z) =$ the standard normal cdf.

The sample size n for which a level α test also has $\beta(\mu') = \beta$ at the alternative value μ' is

$$n = \begin{cases} \left[\dfrac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{for a one-tailed (upper or lower) test} \\ \left[\dfrac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'}\right]^2 & \text{for a two-tailed test (an approximate solution)} \end{cases}$$

Chapter 9 Stuff

Comparing $\mu_1 - \mu_2$

The properties of $\bar{X}_1 - \bar{X}_2$: difference of sample

1. $E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2) = \mu_1 - \mu_2$ unbiased

2. $V(\bar{x}_1 - \bar{x}_2)$: $\bar{x}_1$ and $\bar{x}_2$ are indepedent $V(\bar{x}_1) + V(\bar{x}_2) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}$

$\bar{x}_1$ and $\bar{x}_2$ independent because 2 samples are independent

$\sqrt{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

3. $\bar{x}_1, \bar{x}_2$ both normal $\Rightarrow \bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

4. $P\left(-Z_{\frac{\alpha}{2}} \le \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \le Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$

$P\left(\bar{x}_1 - \bar{x}_2 - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \underbrace{\mu_1 - \mu_2}_{\text{(population) mean difference}} \le (\bar{x}_1 - \bar{x}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 1 - \alpha$

The hypothesis:

$H_0 : \mu_1 - \mu_2 = c.$

Population	**Sample**	σ^2	**Equal** σ^2	**Test statistic**	**CI**
-	Large	Known	-	$\frac{(\bar{X}_1-\bar{X}_2)-c}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$\left[(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$
Normal	-				
-	Large	Unknown	-	$\frac{(\bar{X}_1-\bar{X}_2)-c}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim N(0,1)$	$\left[(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right]$
Normal	Small	Unknown	Equal	$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ $\frac{(\bar{X}_1-\bar{X}_2)-c}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$	$\left[(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}^{(n_1+n_2-2)} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]$
			Not Equal	$v = \left\lfloor \frac{(s_1^2/n_1+s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}} \right\rfloor$ $\frac{(\bar{X}_1-\bar{X}_2)-c}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t(v)$	$\left[(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}^{(v)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right]$

The hypothesis:

(population) mean difference

$H_0 : \mu_1 - \mu_2 = c.$

$H_a : \mu_1 - \mu_2 > c$ **or** $H_a : \mu_1 - \mu_2 < c$ **or** $H_a : \mu_1 - \mu_2 \neq c$

TS: $z = \frac{\bar{X}_1 - \bar{X}_2 - c}{\sqrt{\sigma^2/n_1 + \sigma^2/n_2}}$

RR:

	$H_a : \mu_1 - \mu_2 > c$	$H_a : \mu_1 - \mu_2 < c$	$H_a : \mu_1 - \mu_2 \neq c$
RR	(z_α, ∞)	$(-\infty, -z_\alpha)$	$(-\infty, -z_{\frac{\alpha}{2}}) \cup (z_{\frac{\alpha}{2}}, +\infty)$

We reject H_0 when TS z lies in the RR, or P-value is less than or equal to α.

$\lfloor \; n_1 - 1 \quad , \quad n_2 - 1 \; \rfloor$ Floor / round down

The hypothesis:

$H_0 : \mu_1 - \mu_2 = c.$

$H_a : \mu_1 - \mu_2 > c$ **or** $H_a : \mu_1 - \mu_2 < c$ **or** $H_a : \mu_1 - \mu_2 \neq c$

TS:

$$T = \frac{\bar{X}_1 - \bar{X}_2 - c}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \sim t(v).$$

RR:

	$H_a : \mu_1 - \mu_2 > c$	$H_a : \mu_1 - \mu_2 < c$	$H_a : \mu_1 - \mu_2 \neq c$
RR	$(t_\alpha(v), +\infty)$	$(-\infty, -t_\alpha(v))$	$(-\infty, -t_{\frac{\alpha}{2}}(v)) \cup (t_{\frac{\alpha}{2}}(v), +\infty)$