[redacted]: Introduction to Hypothesis Testing

A manufacturer has been producing shearing pins to be used under certain stress conditions. It is known (assumed)

$$\text{life length of of a pin (in hours)} \sim N(\mu_0 = 100, 9).$$

Suppose a new manufacturing process is introduced to extend the life of these pins.

$$X = \text{life length of a pin by the new process} \sim N(\mu, 9)$$

We hope $\mu > 100$.

(C) We can test the "null hypothesis" $H_0 : \mu = 100$
versus the "alternative hypothesis" $H_1 : \mu > 100$

We collect a random sample $X_1, \ldots, X_n$ from the distribution of the new process.
Based on the data, we have two possible conclusions:

"Reject H_0" : Conclude life length increases ↑
"Accept H_0" : Conclude life length has not changed

Remark: The decision rules that we will derive assign "Reject H_0" as a *strong* conclusion while "Accept H_0" is a *weak conclusion* and is more propoperly stated as "Fail to Reject H_0."

There are two types of *errors* that can occur:

		True State of Nature	
		H_0 true	H_0 false
Statistician's	H_0 true	OK	II
Conclusion	H_0 false	I	OK

A **type I error** occurs when we *reject* a true hypothesis.

A **type II error** occurs when we *accept* a false hypothesis.

For **(C)**, type I error: occurs when we reject a true hypothesis Reject H_0 when H_0 is true
type II error: occurs when we accept a false hypothesis

Suppose instead, a cheaper process is introduced which may decrease the life of these pins.

(B) We can test
$H_0 : \mu = 100$
vs. $H_1 : \mu < 100$
type I error: Conclude ↓ when it really has not gone down
type II error: Conclude not changed when really it has gone down

Or suppose we simply want to know if the new process alters the life length of these pins.

(A) We can test
$H_0 : \mu = 100$
vs. $H_1 : \mu \neq 100$
type I error: Conclude changed when really it has not
type II error: Conclude not changed when really changed

Remark: When possible, the type I error is assigned the more *serious* error.

[redacted]: Paired versus Independent Samples

(1) We are interested in the mean time to parallel park car A versus car B. 14 people volunteer to park the cars.

Suppose all 14 people park both cars A and B (in some randomized order). Label the people $1, \ldots, 14$. Let

$$
\begin{aligned}
X_{1i} &= \text{the time it took for person } i \text{ to park car A}, \; i = 1, \ldots, 14 \\
X_{2i} &= \text{the time it took for person } i \text{ to park car B}, \; i = 1, \ldots, 14
\end{aligned}
$$

Person $i \Rightarrow X_{1i}, X_{2i} \Rightarrow$ dependent, correlated

$\Rightarrow$ PAIRED

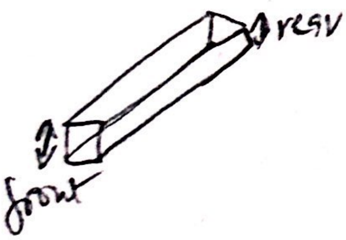
Suppose 6 of the 14 people park car A and and the other 8 park car B. Label those that parked car A as $1, \ldots, 6$ and those that parked car B as $1, \ldots, 8$. Let

$$
\begin{aligned}
X_{1i} &= \text{the time it took for person } i \text{ to park car A}, \; i = 1, \ldots, 6 \\
X_{2j} &= \text{the time it took for person } j \text{ to park car B}, \; j = 1, \ldots, 8
\end{aligned}
$$

different people park car A vs car B

$\Rightarrow$ INDEPENDENT

(2) A machine is used to form lengths of wood with a particular thickness (μ_0). We want to know if the thicknesses at each end of a wood piece are the same.



Front, rear measured on same piece of wood

$\Rightarrow$ PAIRED

(3) We are interested in the comparison of caffeine content between two kinds of instant coffee: spray-dried and freeze-dried. Data are collected from 8 brands of instant coffee, 4 of which are spray-dried and 4 of which are freeze-dried.

8 brands are all different

$\Rightarrow$ INDEPENDENT

Slide Pg 11

Standard Errors

$$\text{s.e.}\{b_1\} = \sqrt{\frac{\text{MSE}}{S_{XX}}} \; ; \; \text{s.e.}\{b_0\} = \sqrt{\text{MSE}\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{XX}}\right]}$$

$$\text{s.e.}\{\hat{Y}|_{x=x_0}\} = \sqrt{\text{MSE}\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}\right]}$$

$100(1-\alpha)\%$ Confidence Intervals

For β_1 two-sided: $b_1 \pm t_{\alpha/2,\,n-2}\,\text{s.e.}\{b_1\}$
upperbound: $(-\infty\,,\; b_1 + t_{\alpha,\,n-2}\,\text{s.e.}\{b_1\})$
lowerbound: $(b_1 - t_{\alpha,\,n-2}\,\text{s.e.}\{b_1\}\,,\; \infty)$

For β_0 two-sided: $b_0 \pm t_{\alpha/2,\,n-2}\,\text{s.e.}\{b_0\}$
upperbound: $(-\infty\,,\; b_0 + t_{\alpha,\,n-2}\,\text{s.e.}\{b_0\})$
lowerbound: $(b_0 - t_{\alpha,\,n-2}\,\text{s.e.}\{b_0\}\,,\; \infty)$

For $E[Y|x = x_0]$ two-sided: $\hat{y}|_{x=x_0} \pm t_{\alpha/2,\,n-2}\,\text{s.e.}\{\hat{Y}|_{x=x_0}\}$
upperbound: $\left(-\infty\,,\; \hat{y}|_{x=x_0} + t_{\alpha,\,n-2}\,\text{s.e.}\{\hat{Y}|_{x=x_0}\}\right)$
lowerbound: $\left(\hat{y}|_{x=x_0} - t_{\alpha,\,n-2}\,\text{s.e.}\{\hat{Y}|_{x=x_0}\}\,,\; \infty\right)$

Hypothesis Tests: Use t-tests with degrees of freedom $v = n - 2$.

	Slope	Y-intercept
Parameter of Interest θ	β_1	β_0
Point estimator $\hat{\theta}$	b_1	b_0
Standard error s.e.$\{\hat{\theta}\}$	s.e.$\{b_1\}$	s.e.$\{b_0\}$

$100(1-\alpha)\%$ Prediction Intervals

Define the *prediction error* for a new response observation Y at $x = x_0$ to be

$$\text{p.e.}\{Y|_{x=x_0}\} = \sqrt{\text{MSE} + \left(\text{s.e.}\{\hat{Y}|_{x=x_0}\}\right)^2}.$$

Then the corresponding two-sided prediction interval is:

$$\hat{y}|_{x=x_0} \pm t_{\alpha/2,\,n-2}\,\text{p.e.}\{Y|_{x=x_0}\}.$$

Slide pg 18

[redacted]: Confidence Intervals, Hypothesis Tests, and Prediction

Notation: $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $S_{XX} = \sum_{i=1}^{n} x_i^2 - \frac{(\sum x_i)^2}{n}$, $S_{XY} = \sum_{i=1}^{n} x_i Y_i - \frac{(\sum x_i)(\sum Y_i)}{n}$

Point estimator for β_1: $\quad b_1 = \frac{S_{XY}}{S_{XX}} = \sum_{i=1}^{n} c_i Y_i$, where $c_i = \frac{x_i - \overline{x}}{S_{XX}}$ (Note: $\sum_{i=1}^{n} c_i = 0$)

Point estimator for β_0: $\quad b_0 = \overline{Y} - b_1\overline{x} = \frac{\sum Y_i}{n} - b_1\frac{\sum x_i}{n}$

Point estimator for $E[Y|x = x_0]$: $\quad \hat{Y}|_{x=x_0} = b_0 + b_1 x_0 = \overline{Y} + b_1(x_0 - \overline{x})$

Sampling Distributions

$$b_1 \sim N\left(\beta_1 , \frac{\sigma^2}{S_{XX}}\right) \quad ; \quad b_0 \sim N\left(\beta_0 , \sigma^2\left[\frac{1}{n} + \frac{\overline{x}^2}{S_{XX}}\right]\right)$$

$$\hat{Y}|_{x=x_0} \sim N\left(\beta_0 + \beta_1 x_0 , \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \overline{x})^2}{S_{XX}}\right]\right)$$

Note: (1)
$$\begin{aligned} \text{Var}(b_1) &= \text{Var}(\sum_{i=1}^{n} c_i Y_i) \\ &= \sum_{i=1}^{n} c_i^2 \text{Var}(Y_i) = \sigma^2 \sum_{i=1}^{n} c_i^2 = \frac{\sigma^2}{S_{XX}} \end{aligned}$$

(2)
$$\begin{aligned} \text{Cov}(\overline{Y}, b_1) &= \text{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i , \sum_{i=1}^{n} c_i Y_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} c_i \,\text{Cov}(Y_i , Y_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} c_i \,\text{Var}(Y_i) = \frac{\sigma^2}{n}\sum_{i=1}^{n} c_i = 0 \end{aligned}$$

(3)
$$\begin{aligned} \text{Var}(b_0) &= \text{Var}(\overline{Y}) + (-1)^2\text{Var}(b_1\overline{x}) - 2\text{Cov}(\overline{Y}, b_1\overline{x}) \\ &= \text{Var}(\overline{Y}) + \overline{x}^2\text{Var}(b_1) - 2\overline{x}\text{Cov}(\overline{Y}, b_1) \\ &= \frac{\sigma^2}{n} + \overline{x}^2\frac{\sigma^2}{S_{XX}} \end{aligned}$$

(4)
$$\begin{aligned} \text{Var}(\hat{Y}|_{x=x_0}) &= \text{Var}(\overline{Y}) + (x_0 - \overline{x})^2\text{Var}(b_1) - 2(x_0 - \overline{x})\text{Cov}(\overline{Y}, b_1) \\ &= \frac{\sigma^2}{n} + (x_0 - \overline{x})^2\frac{\sigma^2}{S_{XX}} \end{aligned}$$

Pg 18

[redacted] Introduction to Hypothesis Testing

A manufacturer has been producing shearing pins to be used under certain stress conditions. It is known (assumed)

$$\text{life length of of a pin (in hours)} \sim N(\mu_0 = 100, \underbrace{9}_{\sigma^2}). \quad \text{known}$$

Suppose a $\boxed{\text{new}}$ manufacturing process is introduced to extend the life of these pins.

$$X = \text{life length of a pin by the new process} \sim N(\mu, 9) \quad \text{unknown}$$

We hope $\mu > 100$.

Form (C) We can test the "null hypothesis" $H_0 : \mu = 100$
versus the "alternative hypothesis" $H_1 : \mu > 100$

We collect a random sample $X_1, \ldots, X_n$ from the distribution of the new process.
Based on the data, we have two possible conclusions:

"Reject H_0" : Conclude life length ↑
"Accept H_0" : Conclude life length has not changed

Remark: The decision rules that we will derive assign "Reject H_0" as a *strong* conclusion while "Accept H_0" is a *weak conclusion* and is more propoperly stated as "Fail to Reject H_0."

There are two types of *errors* that can occur:

		True State of Nature	
		H_0 true	H_0 false
us Statistician's	H_0 true	OK	II
Conclusion	H_0 false	I	OK

Form

A **type I error** occurs when we *reject* a true hypothesis. Reject H_0 when H_0 is true

A **type II error** occurs when we *accept* a false hypothesis. Accept H_0 when H_0 is false

For (C), type I error: Conclude life length ↑ when it really has not
More serious
type II error: Conclude not changed, when it really has gone up ↑

Suppose instead, a cheaper process is introduced which may decrease the life of these pins.

Form (B) We can test

SWAP $H_0 : \mu \leq 0$
vs $H_1 : \mu > 100$

$H_0 : \mu = 100$
vs. $H_1 : \mu < 100$ ← decrease

type I error: Conclude ↓ when it really has not
More serious — type II error: Conclude not changed, when really it has gone down

Or suppose we simply want to know if the new process alters the life length of these pins.

Form (A) We can test

$H_0 : \mu = 100$
vs. $H_1 : \mu \neq 100$

type I error: Conclude change, when it really has not
type II error: Conclude ~~not changed~~, when really changed

Remark: When possible, the type I error is assigned the more *serious* error.