Chapter 7

# MLR Model Evaluation

Chanwoo Yoo, Division of Advanced Engineering,
Korea National Open University

# Contents

# 1. Three Types of Hypotheses

# 1. Three Types of Hypotheses

- A hypothesis test for testing that one slope parameter is 0

- A hypothesis test for testing that all of the slope parameters are 0

- A hypothesis test for testing that a subset — more than one, but not all — of the slope parameters are 0

# 2. Example: Heart attacks in Rabbits

- Does the mean size of the infarcted area differ among the three treatment groups — no cooling, early cooling, and late cooling — when controlling for the size of the region at risk for infarction?

# 2. Example: Heart attacks in Rabbits

- Rabbits whose hearts were cooled to 6° C within 5 minutes of the blocked artery ("early cooling")

- Rabbits whose hearts were cooled to 6° C within 25 minutes of the blocked artery ("late cooling")

- Rabbits whose hearts were not cooled at all ("no cooling")

# 3. Data: Cool Hearts

- Data: Cool Hearts

  - $y_i$ (Inf.): the size of the infarcted area (in grams) of rabbit $i$

  - $x_{i1}$ (Area): the size of the region at risk (in grams) of rabbit $i$

  - $x_{i2}$ (X2): 1 if early cooling of rabbit $i$, 0 if not

  - $x_{i3}$ (X3): 1 if late cooling of rabbit $i$, 0 if not

  - $y_i = (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}) + \epsilon_i$

# 4. Code: Cool Hearts

```
coolhearts <- read.table("coolhearts.txt", header=T)
attach(coolhearts)

model.1 <- lm(Inf. ~ Area + X2 + X3)
summary(model.1)
anova(model.1)
```

# 5. Results: Cool Hearts

```
> summary(model.1)
...
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.13454    0.10402  -1.293 0.206459
Area         0.61265    0.10705   5.723 3.87e-06 ***
X2          -0.24348    0.06229  -3.909 0.000536 ***
X3          -0.06566    0.06507  -1.009 0.321602
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

한국방송통신대학교 프라임칼리지

# 5. Results: Cool Hearts

```
> summary(model.1)

...
Residual standard error: 0.1395 on 28 degrees of freedom
Multiple R-squared:  0.6377,  Adjusted R-squared:  0.5989
F-statistic: 16.43 on 3 and 28 DF,  p-value: 2.363e-06
```

# 5. Results: Cool Hearts

```
> anova(model.1)
Analysis of Variance Table

Response: Inf.
          Df  Sum Sq Mean Sq F value     Pr(>F)
Area       1 0.62492 0.62492 32.1115 4.504e-06 ***
X2         1 0.31453 0.31453 16.1622  0.000398 ***
X3         1 0.01981 0.01981  1.0181  0.321602
Residuals 28 0.54491 0.01946
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
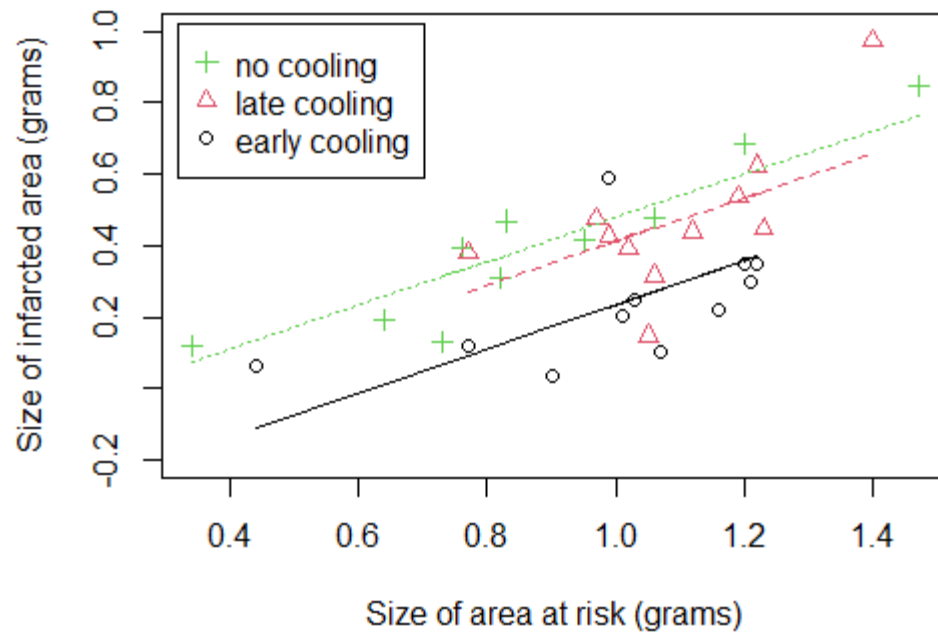
# 6. Code: Cool Hearts

```
plot(Area, Inf., type="n", ylim=c(-0.2, 1),
     xlab="Size of area at risk (grams)",
     ylab="Size of infarcted area (grams)")
for (i in 1:32) points(Area[i], Inf.[i], pch=Group[i],
col=Group[i])
for (i in 1:3) lines(Area[Group==i], fitted(model.1)[Group==i],
lty=i, col=i)
legend("topleft", legend=c("no cooling",
                           "late cooling",
                           "early cooling"),
       col=3:1, pch=3:1, inset=0.02)
```
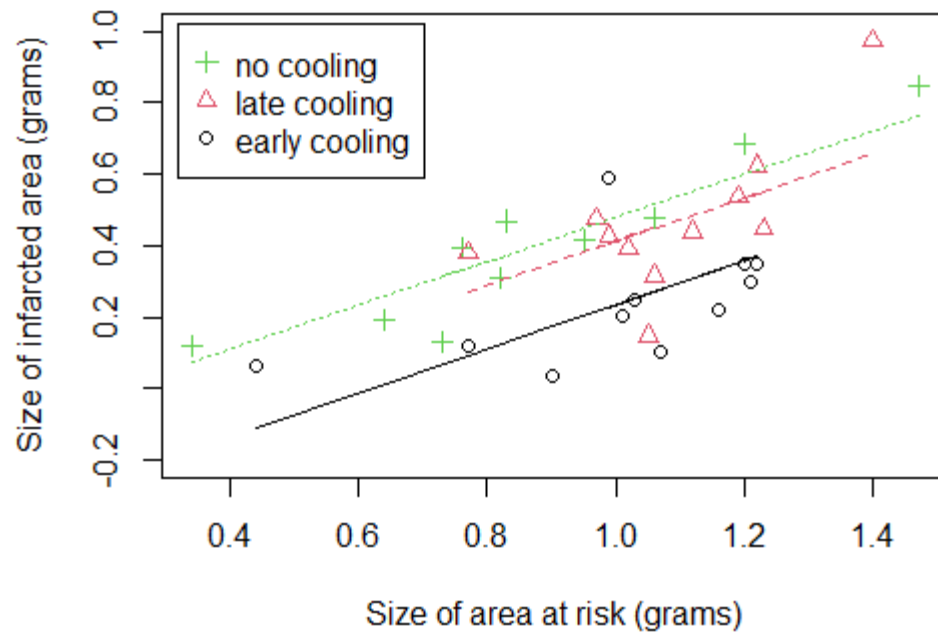
# 7. Results: Cool Hearts



- As the size of the area at risk increases, the size of the infarcted area also tends to increase.

# 7. Results: Cool Hearts
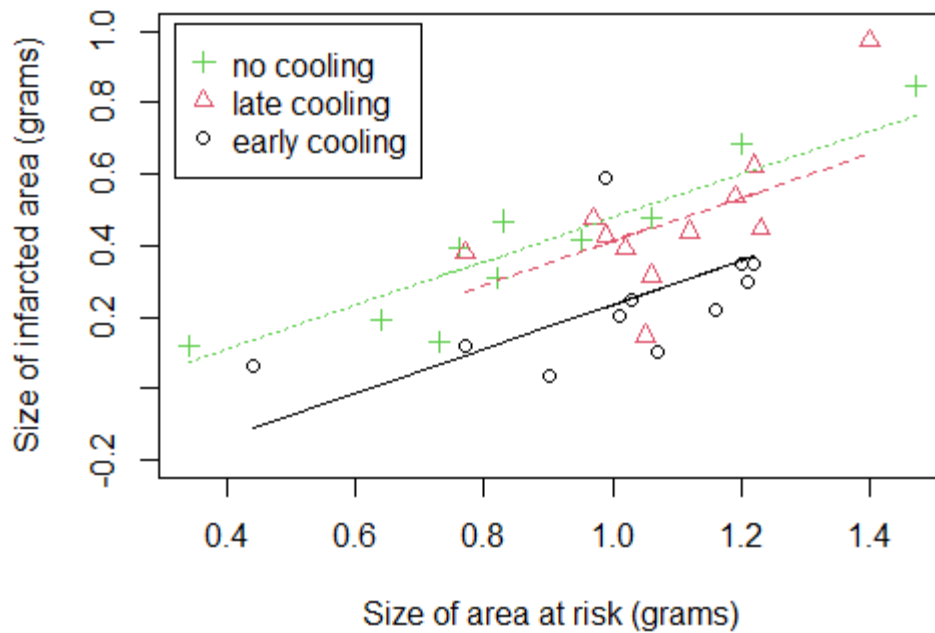


- The plot also suggests that for this sample of 32 rabbits with a given size of area at risk, the average size of the infarcted area differs for the three experimental groups.

# 7. Results: Cool Hearts



- We want to be able to answer the research question for the whole population of rabbits.

# 8. Statistical Procedure

- To translate a research question into a statistical procedure,

  - formulate a multiple regression model

  - determine how the model helps answer the research question

  - check the model

  - performe a hypothesis test (or calculating a confidence interval)

# 9. Research Question I

- Is a regression model containing at least one predictor useful in predicting the size of the infarct?

  - $H_0: \beta_1 = \beta_2 = \beta_3 = 0$

  - $H_A$: At least one $\beta_i \neq 0$ (for $i = 1,2,3$)

# 10. Research Question II

- Is the size of the infarct significantly (linearly) related to the area of the region at risk?

  - $H_0: \beta_1 = 0$

  - $H_A: \beta_1 \neq 0$

# 11. Research Question III

- Is the size of the infarct area significantly (linearly) related to the type of treatment after controlling for the size of the region at risk for infarction?

  - $H_0: \beta_2 = \beta_3 = 0$

  - $H_A$: At least one $\beta_i \neq 0$ (for $i = 2,3$)

# 2. The General Linear F-Test

# 1. The General Linear F-Test

- The "general linear F-test" involves three basic steps

  - Define a larger full model.

  - Define a smaller reduced model.

  - Use an F-statistic to decide whether or not to reject the smaller reduced model in favor of the larger full model.

# 2. The Full Model

- The "full model", which is also sometimes referred to as the "unrestricted model," is the model thought to be most appropriate for the data. For simple linear regression, the full model is

$$y_i = (\beta_0 + \beta_1 x_{i1}) + \epsilon_i$$

한국방송통신대학교
프라임칼리지

# 2. The Full Model
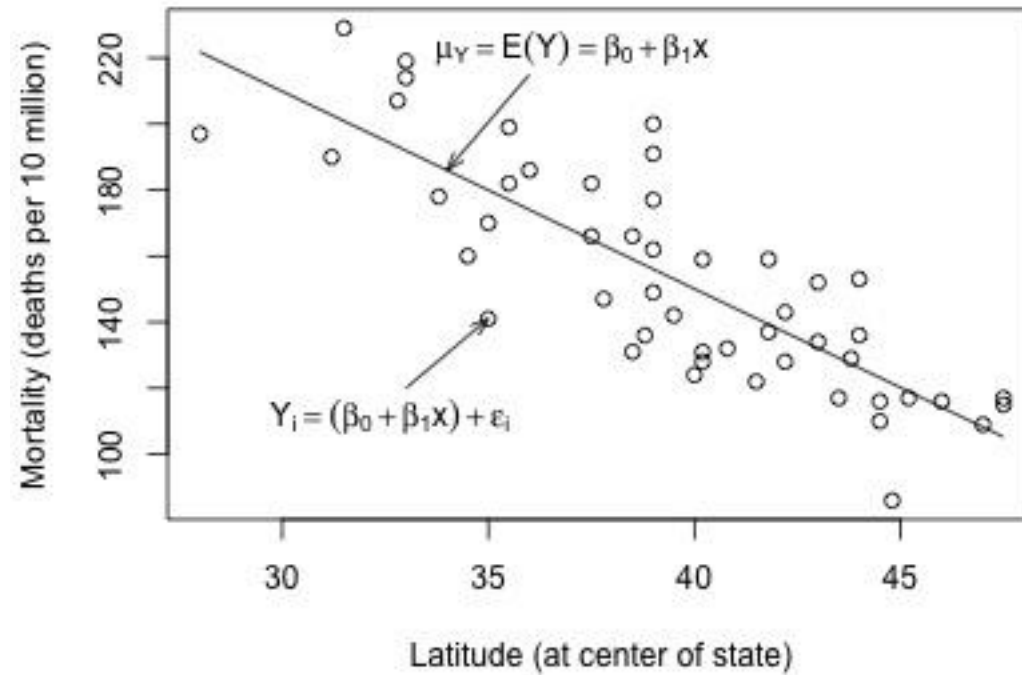
# 3. The Reduced Model

- The "reduced model," which is sometimes also referred to as the "restricted model," is the model described by the null hypothesis $H_0$. For simple linear regression, a common null hypothesis is $H_0 : \beta_1 = 0$. In this case, the reduced model is obtained by "zeroing-out" the slope $\beta_1$ that appears in the full model. That is, the reduced model is

$$y_i = \beta_0 + \epsilon_i$$

# 3. The Reduced Model
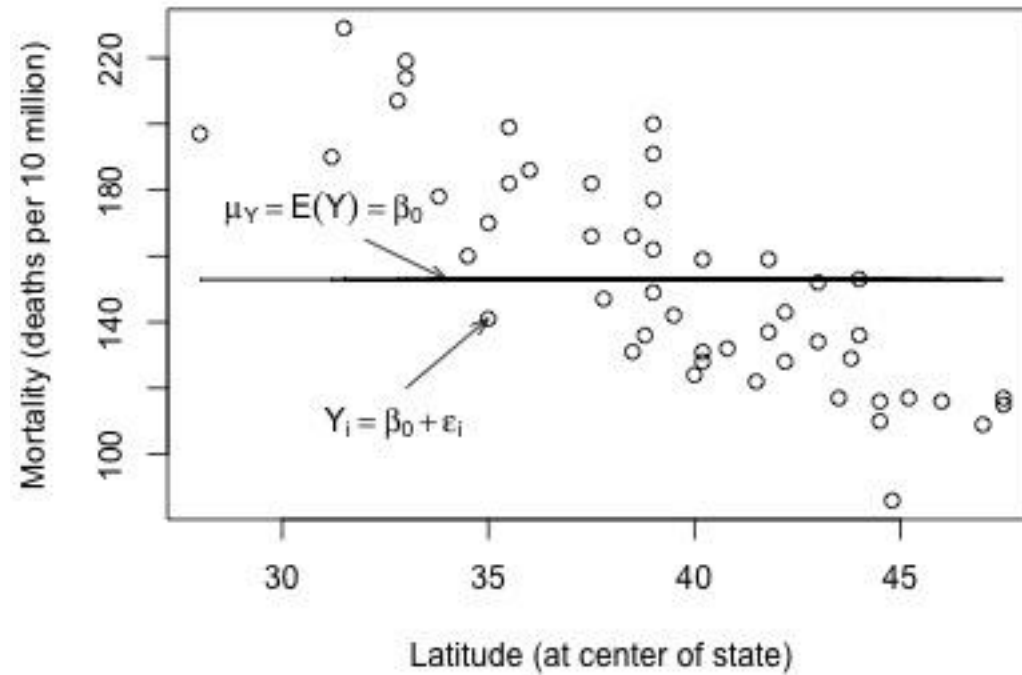
# 4. F-Statistic Test

- How do we decide if the reduced model or the full model does a better job of describing the trend in the data when it can't be determined by simply looking at a plot? What we need to do is to quantify how much error remains after fitting each of the two models to our data. That is, we take the general linear F-test approach.

# 4. F-Statistic Test

- "Fit the full model" to the data.

  - Obtain the least squares estimates of $\beta_0$ and $\beta_1$.

  - Determine the error sum of squares, which we denote "SSE(F)."

- "Fit the reduced model" to the data.

  - Obtain the least squares estimate of $\beta_0$.

  - Determine the error sum of squares, which we denote "SSE(R)."

# 4. F-Statistic Test

- The general linear F-statitstic

  - $F^* = \left(\frac{SSE(R) - SSE(F)}{df_R - df_F}\right) \div \left(\frac{SSE(F)}{df_F}\right)$

  - $df_R, df_F$ : The degrees of freedom

- We use the general linear F-statistic to decide whether or not

  - to reject the null hypothesis $H_0$ : The reduced model

  - in favor of the alternative hypothesis $H_A$ : The full model

## 5. The test applied to the simple linear regression model

- $H_0: y_i = \beta_0 + \epsilon_i$

- $H_A: y_i = (\beta_0 + \beta_1 x_{i1}) + \epsilon_i$

## 5. The test applied to the simple linear regression model

- $F^* = \left(\frac{SSE(R) - SSE(F)}{df_R - df_F}\right) \div \left(\frac{SSE(F)}{df_F}\right)$

  - $df_R = n - 1, df_F = n - 2$

  - $SSE(R) = SSTO, SSE(F) = SSE$

- $F^* = \left(\frac{SSE(R) - SSE(F)}{df_R - df_F}\right) \div \left(\frac{SSE(F)}{df_F}\right) = \left(\frac{SSTO - SSE}{(n-1) - (n-2)}\right) \div \left(\frac{SSE}{n-2}\right) = \frac{MSR}{MSE}$

# 6. Example

```
skincancer <- read.table("skincancer.txt", header=T)
attach(skincancer)
model <- lm(Mort ~ Lat)

summary(model)
anova(model)
```

# 6. Example

```
> anova(model)
Analysis of Variance Table

Response: Mort
            Df Sum Sq Mean Sq F value      Pr(>F)
Lat          1  36464   36464  99.797 3.309e-13 ***
Residuals 47  17173     365
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$MSR$

$SSE$   $MSE$

$SSTO = 36464 + 17173 = 53637$

$SSE = 17173$

$MSR = 36464$

$MSE = 365$

# 6. Example

- $MSR = 36464$

- $MSE = 365$

- $F^* = \dfrac{MSR}{MSE} = \dfrac{36464}{365} = 99.9$

  - 1 numerator degree of freedom and 47 denominator degree of freedom

# 6. Example

```
> pf(99.9, 1, 47, lower.tail = FALSE)
[1] 3.254756e-13
```

# 6. Example

```
> summary(model)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 389.1894    23.8123   16.34  < 2e-16 ***
Lat          -5.9776     0.5984   -9.99 3.31e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.12 on 47 degrees of freedom
Multiple R-squared:  0.6798,  Adjusted R-squared:  0.673
F-statistic:  99.8 on 1 and 47 DF,  p-value: 3.309e-13
```

# 6. Example

```
> anova(model)
Analysis of Variance Table

Response: Mort
          Df Sum Sq Mean Sq F value    Pr(>F)
Lat        1  36464   36464  99.797 3.309e-13 ***
Residuals 47  17173     365
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Next

# Chapter 8
# MLR Model Evaluation II