Level M Regression Models

Lecture 11

Colette Mair

School of Mathematics and Statistics

Hypothesis testing

Introduction

- In the previous week we considered diagnostics for and assumptions about linear regression models. Now we will consider inference for model parameters
- We will focus on hypothesis testing of the regression parameters.

Linear Combinations of Parameters

Linear Combinations of Parameters

- We will develop a general theory for doing inference on linear combinations of parameters.
- For example if we have a simple linear regression

$$y_i = \beta_0 + \beta_1 x + \epsilon_i \implies \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x$$

- We want to predict a future value, at $x = 5$.

Linear Combinations of Parameters

- We are interested in the linear combination:

$$\beta_0 + 5\beta_1$$

- This can be written as:

$$(1 \quad 5)\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \qquad .$$

Linear Combinations of Parameters

- We are interested in the linear combination:

$$\beta_0 + 5\beta_1$$

- This can be written as:

$$(1 \quad 5)\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \phi.$$

Multiple linear model

Multiple linear model

Data: $(y_i, x_{1i}, x_{2i}, \dots, x_{ki}); \quad i = 1, \dots, n$
Model: $\mathbf{Y} = \mathbf{X}\beta + \epsilon$ with $\epsilon = \{\epsilon_1, \dots, \epsilon_n\} \sim N(0, \sigma^2)$
independent.

Multiple linear model

- Say $\mathbf{b}_1^T\beta$ for some given vector $\mathbf{b}_1$. For instance in the previous example $\mathbf{b}_1^T = (1 \quad 5)$,
- Possibly for a set of s linearly independent linear combinations $\mathbf{b}_1^T\beta, \ldots, \mathbf{b}_s^T\beta$, $s \leq p$ where the $\mathbf{b}_i$'s are given vectors (similar to the previous example).
- Here, p is the number of regression coefficients, hence β is a vecotr of length p.

Multiple linear model

It is always possible to create a non-singular transformation from $\beta \leftrightarrow \phi$ where

$$\phi = \begin{pmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ . \\ . \\ . \\ \mathbf{b}_s^T \end{pmatrix} \beta = \mathbf{B}\beta,$$

$$\beta = \mathbf{B}^{-1}\phi$$

Multiple linear model

Data: $(y_i, x_{1i}, x_{2i}, \dots, x_{ki}); \quad i = 1, \dots, n$

Model: $\mathbf{Y} = \mathbf{X}\beta + \epsilon = (\mathbf{X}\mathbf{B}^{-1})\phi + \epsilon$

Multiple linear model

Hence, we can write down the solution for the parameter estimates, based on least-squares, from our earlier results.

$\hat{\phi} =$

Multiple linear model

Hence the least-squares estimates of a set of linear functions of parameters is just the set of linear functions of the least-squares estimates.

Example

Example

Data: $(y_i, x_i); \quad i = 1, ..., n$

Model: $y_i = \alpha + \beta x_i + \epsilon_i$, or in vector matrix notation
$\mathbf{Y} = \mathbf{X}\beta + \epsilon = (\mathbf{X}\mathbf{B}^{-1})\phi + \epsilon$

Suppose we transform

$y_i = \alpha + \beta x_i + \epsilon_i$ (Model 1) to

$y_i = \alpha' + \beta(x_i - \bar{x}) + \epsilon_i$ (Model 2)

Example

Hypothesis Testing

Pivotal function for a linear function of the parameters

$$\frac{(\mathbf{b}^T\hat{\beta} - \mathbf{b}^T\beta)}{\sqrt{\frac{RSS}{n-p}\mathbf{b}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{b}}}$$

is a pivotal function

Pivotal function for a linear function of the parameters

$$\frac{(\mathbf{b}^T\hat{\beta} - \mathbf{b}^T\beta)}{\sqrt{\frac{RSS}{n-p}\mathbf{b}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{b}}} \sim t(n-p),$$

- p is the number of estimate parameters
- n is the sample size
- RSS is the residual sum-of-squares in a linear model.

Estimated standard error

$$ese(\mathbf{b}^T\hat{\beta}) = \sqrt{\frac{RSS}{n-p}\mathbf{b}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{b}}$$

The above result can be used to construct hypothesis tests and interval estimates for model parameters.

Hypothesis Testing

- Suppose we were interested in making inferences about β in a simple linear regression model i.e.

$$y_i = \alpha + \beta x_i + \epsilon_i$$

$\mathbf{b}^T \beta = \beta$ i.e. $\mathbf{b}^T = (0 \quad 1)$

$$\frac{\hat{\beta} - \beta}{\text{e.s.e}(\hat{\beta})} \sim t(n-p)$$

Hypothesis Testing

- Under the null hypothesis:

H$_0$: $\beta = 0$

H$_1$: $\beta \neq 0$

$$\frac{\hat{\beta} - 0}{\text{e.s.e}(\hat{\beta})} = \frac{\hat{\beta}}{\text{e.s.e}(\hat{\beta})} \sim t(n-p)$$

Example

Hypothesis testing for Pregnancy Data

- Data were collected through interest in whether, and if so, in what way the level of protein changes in expectant mothers throughout their pregnancy.
- Observations were taken on 19 healthy women. Each woman was at a different stage of pregnancy, gestation.
- We have seen this example previously for parameter estimation assessing model fit.

Hypothesis testing for Pregnancy Data

Data: $(y_i, x_i) \quad i = 1, \ldots, 19$

Model: $E(Y_i) = \alpha + \beta x_i$

Perform a hypothesis test to test $H_0 : \beta = 0$.

term	estimate	std.error	statistic	p.value
(Intercept)	0.2017	0.0834	2.4200	0.027011
Gestation	0.0228	0.0033	6.9337	0.000002

Hypothesis testing for Pregnancy Data

The hypotheses being tested for the coefficient of β are:

$H_0 : \beta = 0$

$H_1 : \beta \neq 0$

From the regression output

$$\frac{\hat{\beta}}{\text{e.s.e}(\hat{\beta})} = \frac{0.022844}{0.003295} = 6.932929 \sim t(n-p) \text{ under } H_0.$$

Hypothesis testing for Pregnancy Data

Since the p-value for gestation is < 0.001 (and hence < 0.05) the null hypothesis is rejected and we conclude that there is a statistically significant relationship between protein and gestation. The gestational age is a useful predictor of the protein level.

Hypothesis testing for Pregnancy Data

$$\frac{\hat{\beta} - 0}{\text{e.s.e}(\hat{\beta})} = \frac{\hat{\beta}}{\text{e.s.e}(\hat{\beta})} \sim t(n-p)$$

$$\frac{0.0228 - 0}{0.0033} = \frac{0.0228}{0.0033} \sim t(19-2)$$

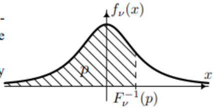
under H_0.

Hypothesis testing for Pregnancy Data

4 Student's t distribution

Inverse $F_\nu^{-1}(p)$ of the cumulative distribution function (quantiles)

The table below contains the quantiles of Student's t distribution with ν degrees of freedom. For $0 < p < 1$ the quantile is the value of x for which $\mathbb{P}\{X \le x\} = p$, where $X \sim \mathrm{t}(\nu)$. Thus $x = F_\nu^{-1}(p)$.
The table only contains the quantiles for $p \ge \frac{1}{2}$. For $p < \frac{1}{2}$ quantiles can be obtained by exploiting the symmetry of the t distribution: $F_\nu^{-1}(p) = -F_\nu^{-1}(1-p)$.



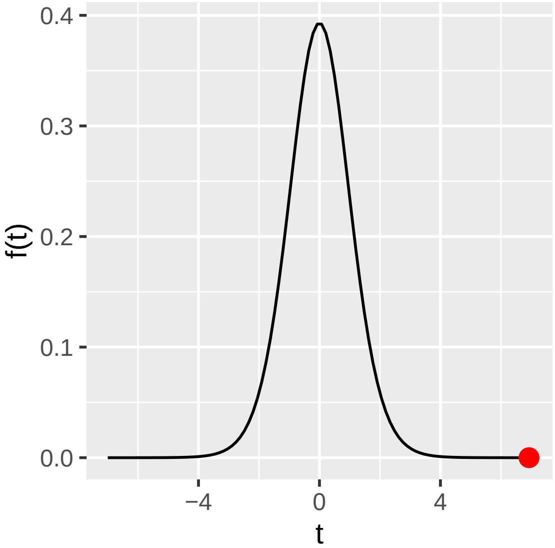
ν \ p	0.6	0.7	0.75	0.8	0.85	0.9	0.95	0.975	0.99	0.995	0.999	0.9995
1	0.3249	0.7265	1.0000	1.3764	1.9626	3.0777	6.3138	12.706	31.821	63.657	318.31	636.62
2	0.2887	0.6172	0.8165	1.0607	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248	22.327	31.599
3	0.2767	0.5844	0.7649	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409	10.215	12.924
4	0.2707	0.5686	0.7407	0.9410	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.2672	0.5594	0.7267	0.9195	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.2648	0.5534	0.7176	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.2632	0.5491	0.7111	0.8960	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.2619	0.5459	0.7064	0.8889	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.2610	0.5435	0.7027	0.8834	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.2602	0.5415	0.6998	0.8791	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.2596	0.5399	0.6974	0.8755	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.2590	0.5386	0.6955	0.8726	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.2586	0.5375	0.6938	0.8702	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.2582	0.5366	0.6924	0.8681	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.2579	0.5357	0.6912	0.8662	1.0735	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.2576	0.5350	0.6901	0.8647	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.2573	0.5344	0.6892	0.8633	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.2571	0.5338	0.6884	0.8620	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.2569	0.5333	0.6876	0.8610	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.2567	0.5329	0.6870	0.8600	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	0.2566	0.5325	0.6864	0.8591	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	0.2564	0.5321	0.6858	0.8583	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	0.2563	0.5317	0.6853	0.8575	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.2562	0.5314	0.6848	0.8569	1.0593	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454

Hypothesis testing for Pregnancy Data

From statistical tables,

- $t(17; 0.975) = 2.1098.$
- the observed test statistic is 6.93
- since $|6.93| > 2.1098$, we have evidence to reject the null hypothesis.

Hypothesis testing for Pregnancy Data



Analysing the ANOVA table

Analysing the ANOVA table

The F statistic value: $\mathsf{MS}_{\text{model}}/\mathsf{MS}_{\text{residuals}}$ provides a test statistic that allows us to test whether there is any evidence that at least one of the model parameters is not zero.

The null hypothesis is

H_0: all p parameters $= 0$,

which will be tested against the alternative that at least one of the parameters is not zero.

Analysing the ANOVA table

Analysing the ANOVA table

If the null hypothesis is true, the statistic has an F(Df_{model}, $Df_{\mathrm{residuals}}$) distribution. This implies that

$$F = \frac{MS_{\mathrm{model}}}{MS_{\mathrm{residuals}}} \sim F(Df_{\mathrm{model}}, Df_{\mathrm{residuals}}).$$

If H_0 is false, we would expect $MS_{\mathrm{residuals}}$ to be smaller than MS_{model} and so large values of F should lead us to reject H_0. i.e. for large values of F the p-value will be small.

Protein in Pregnancy

Looking again at the anova table for protein in pregnancy,

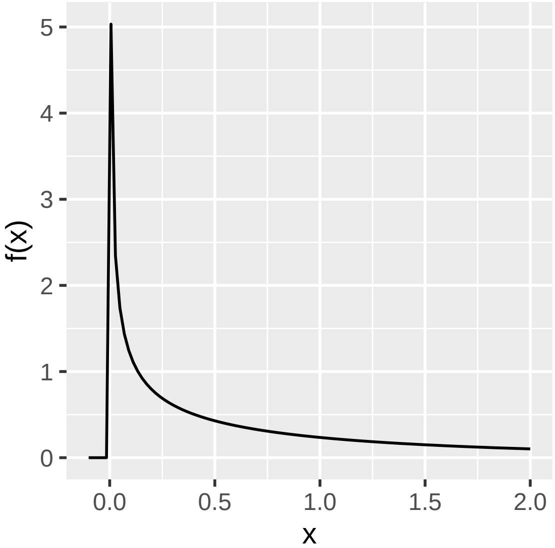
```
## Analysis of Variance Table
##
## Response: Protein
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## Gestation  1 0.63667 0.63667  48.076 2.416e-06 ***
## Residuals 17 0.22513 0.01324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.
```

We can reject the null hypothesis with a p-value of 2.4×10^{-6} suggesting that at least one model parameter is not zero.

Protein in Pregnancy

- The F value is 48.076
- Using an α of 0.05, we have F(0.05, 1,17) $= 4.4513$
- Since 48.076 $>$ 4.4513 we can reject the null hypothesis

Protein in Pregnancy



Protein in Pregnancy