Parameter Estimation Report 1

Keyuan Wang, 13594869

February 15, 2022

1 Exercise 1
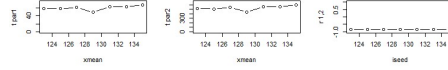

Figure 1: *lmdemoiseed()*

For x$\in$ [123, 135] the t-values for θ_0 and θ_1 are higher than 2, which means the parameters are all significant. The $r_{\theta_0\theta_1}$ remains the same, at a value near -1.0.
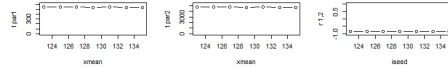

Figure 2: *lmdemoiseed(step=0.01)*

smaller steps leads to much higher t-value for θ_0, others remains same as the previous conclusion.
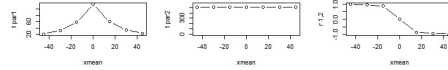

Figure 3: *lmdemorange()*

For x closer to 0, the std error for θ_0 is smaller, t-value is larger. But it does not affect θ_1. All parameters are significant.
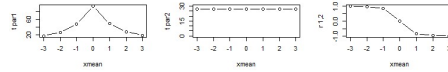

Figure 4: *lmdemorange(range=2,step=0.1)*

By narrowing the range and step size, the t-values are generally smaller for both parameters. And the std errors increased a lot.
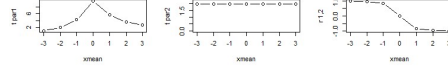
Figure 5: *lmdemorange(range=2,step=0.1, sigma=1)*

The t-value for both parameters are too small, so both parameters are not significant. The std errors are very large.
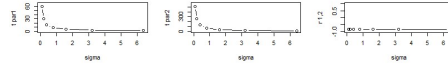

Figure 6: *lmdemosigma()*

With increasing σ value, the t-value for both θ_0 and θ_1 decreases rapidly, but $r_{\theta_0 \theta_1}$ keeps unchanged. The reason is that large noise covered the relationship between y and x.

	std error $\hat{\theta}(\hat{\sigma}_i)$	t-value	$\Pr(> |t|)$	Estimate	$r_{\theta_0 \theta_1}$
olmdemo1$sum1	0.035	58.55	***	2.031	-0.86
	0.002	485.65	***	0.998	
olmdemo1$sum2	0.018	112.2	***	1.992	0
	0.002	485.7	***	0.998	
olmdemo2$sum1	0.026	77.76	***	2.001	-0.71
	0.001	796.56	***	1.000	
olmdemo2$sum2	0.018	109.7	***	1,995	0
	0.001	796.6	***	1.000	

Table 1: four designs

By reading the table, it is obvious that the two experiment designs with $x \in [-15, 15]$ give best fitting results, with higher t-value for both parameters, zero correlation between parameters, and a lower std error, especially for intercept parameter θ_0. In principal, large step size should result in big error, but with 15 times of repeat, the errors seem to be effectively reduced, so the errors in olmdemo2 are generally smaller than in olmdemo1, and the t-values are generally higher, especially for slope parameter θ_1.

From the **Equation 2.18** in the lecture notes we know that for straight line fitting, the minimum variance bound (MVB) only depends on the choice of independent variables x_i alone. With a symmetric design of independent variables, the term $\left(\sum x_i\right)^2 = 0$, resulting in a lower MVB. That is why designs with $x \in [-15, 15]$ have better performance.
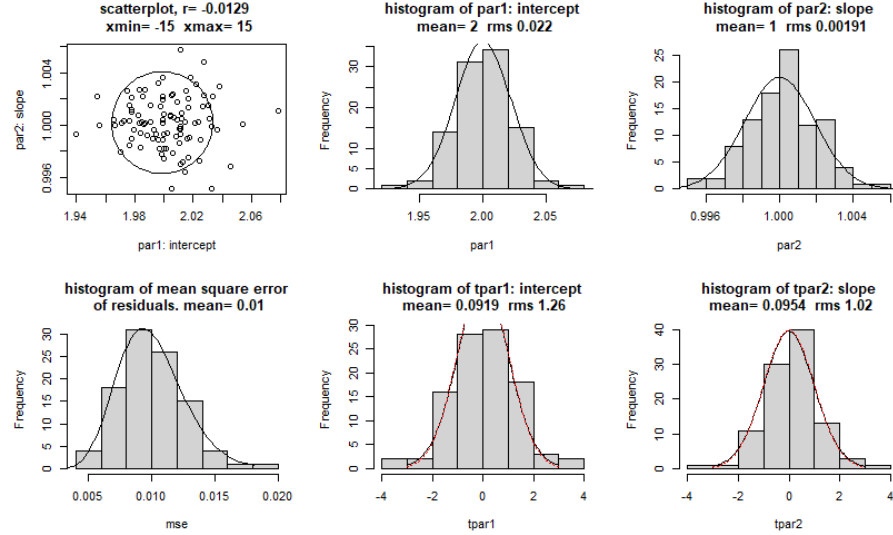
2 Exercise 2

Figure 7: parcorrtdemo(lellipse=T,iseed=12345)

The underlying theoretical distributions of the histograms are all normal distribution. The scatter plot corresponds to the **Equation 2.67** in lecture notes, which is used to calculate the 90% confidence ellipse. The first and the second histograms are distributions of parameter 1 and parameter 2 for multiple simulations, the normal distribution is determined by **Equation 2.57**. The root mean square (rms) value corresponds to the **Equation 1.19** in the lecture notes. For the histogram of mean square error of residuals, the **Equation 2.62** gives it, which is also noted as *residual sum of squares*. The last two histograms are distributions of t-values for parameter 1 and parameter 2, corresponding to the **Equation 2.58** in the lecture notes.

From these plots, we can draw a conclusion that there are no correlation between parameter 1 θ_0 and parameter 2 θ_1, and the most likely value for θ_0 and θ_1 are 2 and 1. However, from the last 2 histograms we can also see that most of the absolute t-values for both parameters are smaller that 2, so we can draw a conclusion that these two parameters are both insignificant.
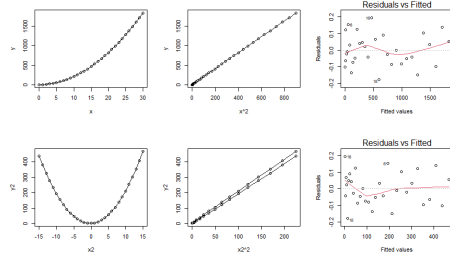
4

3 Exercise 3

Figure 8: mlmdemo()

	Estimate	Std. Error	t-value	Pr($|t|$)	r
	3.047	0.050	61.5	***	$r_{\theta_0\theta_1} = -0.85$
E1\$S1	0.994	0.008	130.0	***	$r_{\theta_0\theta_2} = 0.72$
	2.000	0	8123	***	$r_{\theta_1\theta_2} = -0.97$
	2.985	0.026	113	***	$r_{\theta_0\theta_1} = 0$
E1\$S2	0.998	0.002	507.5	***	$r_{\theta_0\theta_2} = -0.75$
	2.000	0	8123	***	$r_{\theta_1\theta_2} = 0$
	3.007	0.031	97.51	***	$r_{\theta_0\theta_1} = -0.59$
E2\$S1	1.003	0.005	191.42	***	$r_{\theta_0\theta_2} = 0.41$
	2.000	0	11912.03	***	$r_{\theta_1\theta_2} = -0.96$
	3.021	0.031	97.95	***	$r_{\theta_0\theta_1} = 0$
E2\$S2	0.998	0.001	686.66	***	$r_{\theta_0\theta_2} = -0.82$
	2.000	0	11912.03	***	$r_{\theta_1\theta_2} = 0$

Table 2: Diagnostics for four designs

The figure above shows two experiments with different designs. From the residual figure we can see that the residuals oscillate around 0 and have no obvious trend in both experiments, and the t-values are all large so all the parameters in two experiments are significant, so we can conclude that they all have nice fitting results, and the symmetric design is more efficient since the residual curve becomes flat earlier that the non-symmetric design.
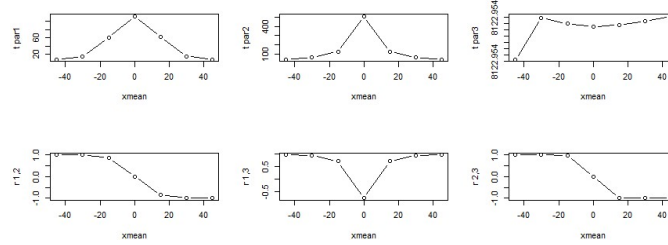

Figure 9: mlmdemorange()

The figure above is similar to Figure 3 in Exercise 1. Just as same as in Exercise 1, the symmetric design always gives the best fitting outcomes: largest t-value and lowest std error, except for parameter 3, the t-value of which reaches peak for negative x values, and std erorr of which does not change during the experiments, we believe the reason is that parameter 3 has the highest order.

From the table we can read that for symmetric experiment design, the correlation between θ_0 (intercept) and θ_1 (I(x)), θ_1 and θ_2 (I(x^2)) are always 0. The reason is that, with a symmetric design, odd and even functions are orthogonal. So within a function, the correlation between terms with odd and even orders are always 0. We believe this argument also holds for higher order functions, i.e. the correlation between first order term and fourth order term is zero, the correlation between third order term and sixth order term is zero, ... etc.
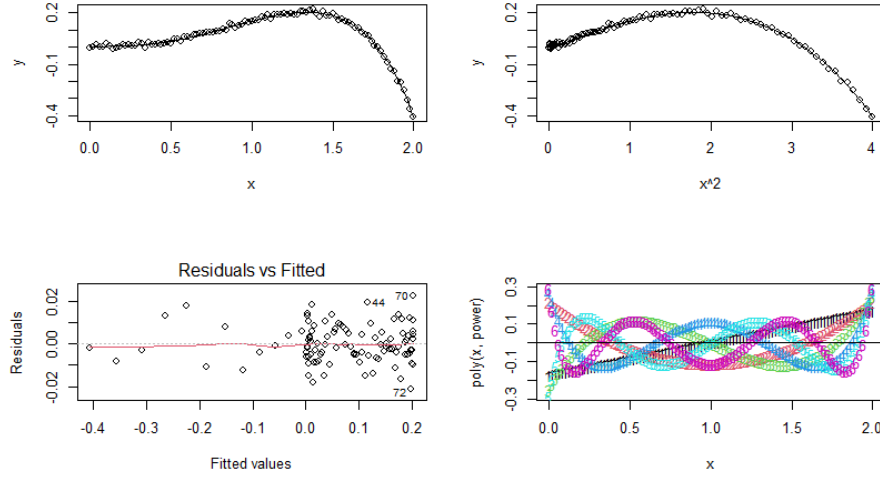
5

4 Exercise 4

Figure 10: Best fitting model: orthogonal polynomial of sixth-order

The formula of my model:

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4 + \theta_5 x^5 + \theta_6 x^6 \tag{1}$$

We select the sixth-order orthogonal polynomial model as the best fit, since it gives the lowest residual standard error and lowest parameter std error, also, all the parameters are significant. For orthogonal polynomials with order higher than 6, the parameters after θ_6 are all insignificant, associated with t-values smaller than 2. The reason why we select orthogonal polynomials instead of ordinary polynomials is that the fitting outcomes for orthogonal polynomials are predictable, that is, by increasing the order, the significance of exiting parameters will not change. But things are different for ordinary polynomials, the parameters used to be significant might become insignificant if increasing the order.

6

5   Exercise 5

The first model given by *powerfitplot()* contains 4 spectra, the second model only contains the first 3 spectra and the third model contains the first 2 spectra. The first model and the second model both have nice fitting outcomes, with low residual errors, but the fourth parameter in the first model has a very low t-value, which means it is insignificant. The third model has bad fitting outcomes with high residual errors comparing with others. So we would say the second model (*powerfitplotomit(select=c(1:4))*) is the best one.