26: INFERENCE FOR SLR

Stat250 S25

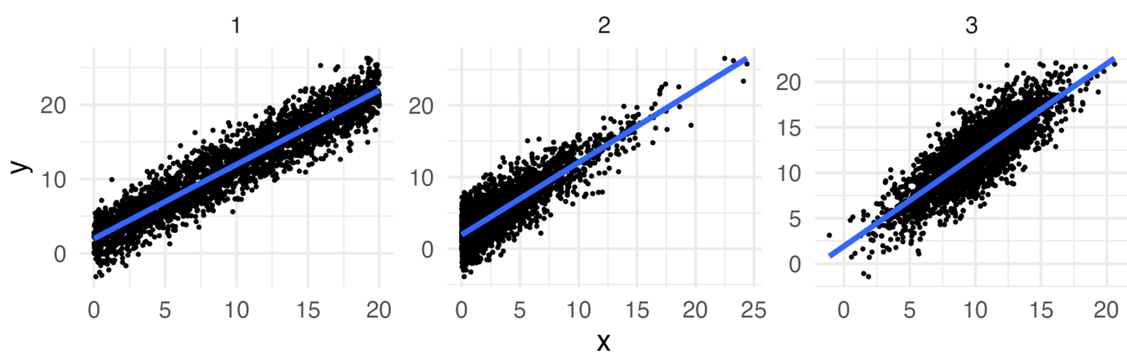
Prof Amanda Luby

1 SLR Model Recap

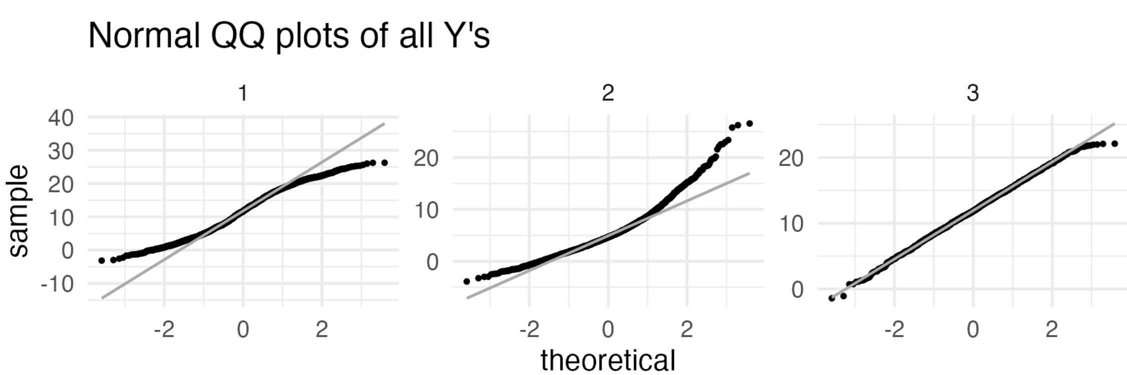
$$Y_i|x_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$$

Key features:

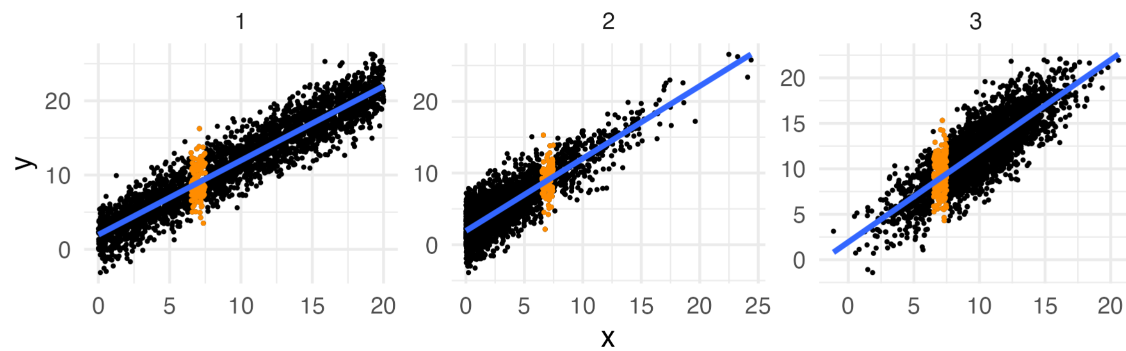
All three of these data sets were generated from the SLR model:



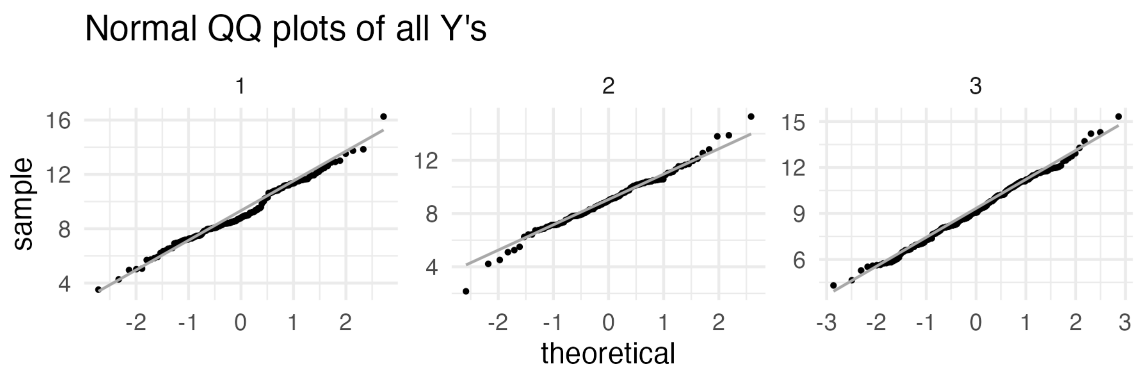
But only data 3 has (marginally) normally distributed responses:



Let's look at the Y values for all cases with $x \approx 7$:



The Y values are (conditionally) normally distributed for with $x \approx 7$



2 Inference for coefficients

Maximum Likelihood Estimators for SLR

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(Y_i - \bar{Y})}{\sum(x_i - \bar{x})^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n}$$

Properties of the MLE's for SLR:

1. $\hat{\beta}_0$ and $\hat{\beta}_1$ are Normally-distributed random variables
2. $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimators
3. $V(\hat{\beta}_1) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$

4. $V(\hat{\beta}_0) = \sigma^2[\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}]$

5. $\hat{\beta}_1, \bar{Y}$ and $\hat{\sigma}^2$ are mutually independent

6. $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-2}$

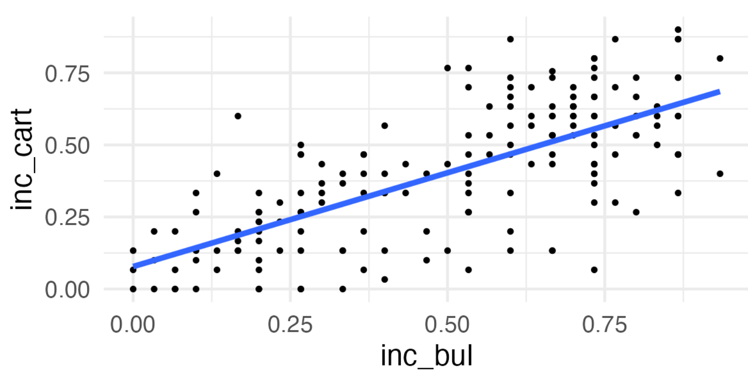
7. $s^2 = \frac{n}{n-2}\hat{\sigma}^2$ is an unbiased estimator for σ^2

Test statistic for β_1

Let $(x_1, Y_1), (x_2, Y_2), ..., (x_n, Y_n)$ be a set of points satisfying $E(Y|X = x) = \beta_0 + \beta_1 x$ and let $S^2 = \frac{1}{n-2}\sum(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2$. Then,

$T = \frac{\hat{\beta}_1 - \beta_1}{S/\sqrt{\sum(x_i - \bar{x})^2}}$

Example: We are interested in the average *inconclusive rate* (how often the firearms examiner cannot come to a definitive conclusion) for bullets compared to cartridge cases. Some results from a study are included below:




```
cb_mod <- lm(inc_cart ~ inc_bul, data = firearms)
summary(cb_mod)
```

```
Call:
lm(formula = inc_cart ~ inc_bul, data = firearms)
```

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.48856 -0.11666  0.01347  0.10022  0.41347 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.07809    0.02811   2.778  0.00609 ** 
inc_bul      0.65064    0.04972  13.087  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1671 on 171 degrees of freedom
Multiple R-squared:  0.5004,	Adjusted R-squared:  0.4975 
F-statistic: 171.3 on 1 and 171 DF,  p-value: < 2.2e-16
```

What is a 95% confidence interval for β_1?

3 Inference for $\mu_{Y|x}$

Example: How do we find a confidence interval for the line?

4 Inference for new data points

Example: If we observe a new examiner who was inconclusive 75% of time time on bullets, what would we predict for their inconclusive rate on cartridge cases?

```
predict(cb_mod, newdata = data.frame(inc_bul = .75))
>         1
> 0.5660712
predict(cb_mod, newdata = data.frame(inc_bul = .75), interval = "confidence")
>         fit       lwr       upr
> 1 0.5660712 0.5312891 0.6008534
predict(cb_mod, newdata = data.frame(inc_bul = .75), interval = "prediction")
>         fit       lwr       upr
> 1 0.5660712 0.2343576 0.8977849
```