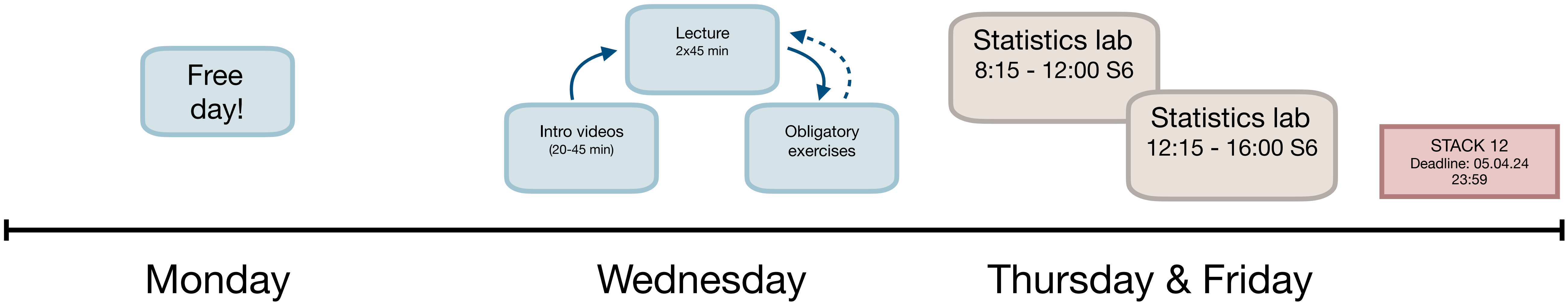
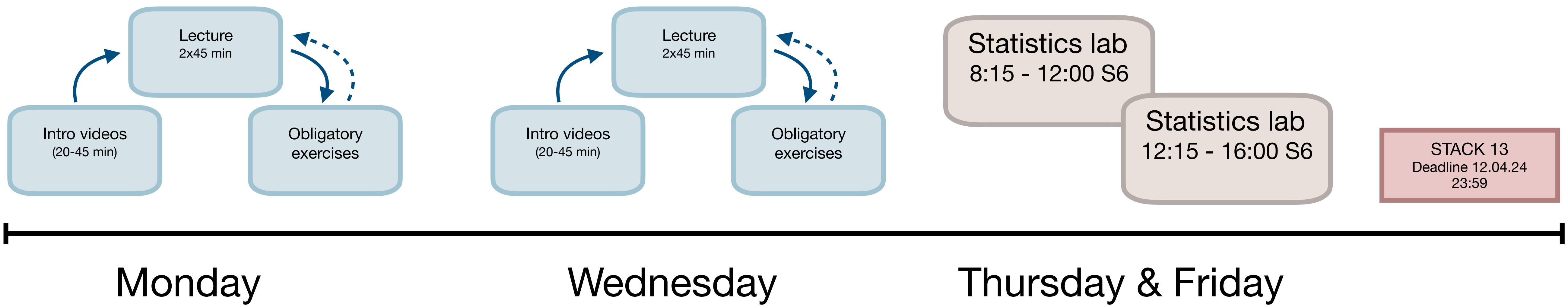
TMA4245 Statistikk

Week 16 - Monday

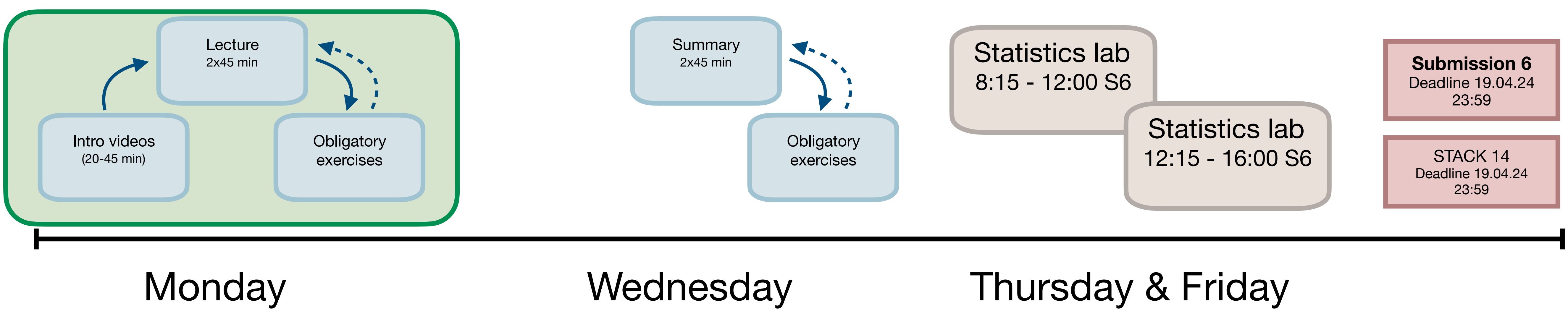
Week 14 (Mon. April 1 - Friday April 5)



Week 15 (Mon. April 8 - Friday April 12)



Week 16 (Mon. April 15 - Friday April 19)



Team based learning: Class structure

Individual quiz

- Find the right answer alone
- Choose the answer that fits best
- No helping material

07:00

Team quiz

- Find the right answer in your teams
- Choose the answer that fits best
- No helping material

10:00

Feedback

Question 1: Which of the following is the LEAST important assumption for simple linear regression?

A $\varepsilon_1, \ldots, \varepsilon_n$ are Gaussian

B $\varepsilon_1, \ldots, \varepsilon_n$ have mean zero

C $\varepsilon_1, \ldots, \varepsilon_n$ are independent

D $(x_1, Y_1), \ldots, (x_n, Y_n)$ have a linear relationship

Question 2: Which of the following is NOT an assumption used in simple linear regression?

A $x_1, \ldots, x_n$ are independent

B $\varepsilon_1, \ldots, \varepsilon_n$ are Gaussian

C $\varepsilon_1, \ldots, \varepsilon_n$ have mean zero

D $(x_1, Y_1), \ldots, (x_n, Y_n)$ have a linear relationship

Question 3: What are the least squares estimators for α and β, the intercept and slope respectively in simple linear regression?

A $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n x_i(x_i - \bar{x})}$

B $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$

C $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (Y_i - (b_0 + b_1 x_i))^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$

D $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$

Question 4: What are the maximum likelihood estimators for α, β, and σ^2?

A $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n x_i(x_i - \bar{x})}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - (b_0 + b_1 x_i))^2$

B $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$, $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n Y_i^2$

C $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$, $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}x_i)^2$

D $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$, $\hat{\sigma}^2 = \sum_{i=1}^n Y_i^2$

Question 5: What is the difference between a prediction interval and a confidence interval in simple linear regression?

A A $(1-\alpha) \times 100\%$ confidence interval has $(1-\alpha) \times 100\%$ chance of containing the prediction μ_0 at x_0, whereas a $(1-\alpha) \times 100\%$ prediction interval has $(1-\alpha) \times 100\%$ chance of containing the response Y_0 at x_0

B A confidence interval is a prediction interval, but a prediction interval is not a confidence interval.

C A confidence interval is a prediction interval, but a prediction interval is not a confidence

D A $(1-\alpha) \times 100\%$ confidence interval has $(1-\alpha) \times 100\%$ chance of containing response Y_0 at x_0, whereas a $(1-\alpha) \times 100\%$ prediction interval has $(1-\alpha) \times 100\%$ chance of containing the prediction μ_0 at x_0

Question 1: Which of the following is the LEAST important assumption for simple linear regression?

A $\varepsilon_1, \dots, \varepsilon_n$ are Gaussian

B $\varepsilon_1, \dots, \varepsilon_n$ have mean zero

C $\varepsilon_1, \dots, \varepsilon_n$ are independent

D $(x_1, Y_1), \dots, (x_n, Y_n)$ have a linear relationship

Question 2: Which of the following is NOT an assumption used in simple linear regression?

A $x_1, \dots, x_n$ are independent

B $\varepsilon_1, \dots, \varepsilon_n$ are Gaussian

C $\varepsilon_1, \dots, \varepsilon_n$ have mean zero

D $(x_1, Y_1), \dots, (x_n, Y_n)$ have a linear relationship

Question 3: What are the least squares estimators for α and β, the intercept and slope respectively in simple linear regression?

A $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n x_i(x_i - \bar{x})}$

B $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$

C $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (Y_i - (b_0 + b_1 x_i))^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$

D $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$

Question 4: What are the maximum likelihood estimators for α, β, and σ^2?

A $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n x_i(x_i - \bar{x})}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - (b_0 + b_1 x_i))^2$

B $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$, $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n Y_i^2$

C $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$, $\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} x_i)^2$

D $\hat{\alpha} = \bar{Y}$, $\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$, $\hat{\sigma}^2 = \sum_{i=1}^n Y_i^2$

Question 5: What is the difference between a prediction interval and a confidence interval in simple linear regression?

A A $(1-\alpha) \times 100\%$ confidence interval has $(1-\alpha) \times 100\%$ chance of containing the prediction μ_0 at x_0, whereas a $(1-\alpha) \times 100\%$ prediction interval has $(1-\alpha) \times 100\%$ chance of containing the response Y_0 at x_0

B A confidence interval is a prediction interval, but a prediction interval is not a confidence interval.

C A confidence interval is a prediction interval, but a prediction interval is not a confidence

D A $(1-\alpha) \times 100\%$ confidence interval has $(1-\alpha) \times 100\%$ chance of containing response Y_0 at x_0, whereas a $(1-\alpha) \times 100\%$ prediction interval has $(1-\alpha) \times 100\%$ chance of containing the prediction μ_0 at x_0

Brief review: Confidence intervals for μ_0

Idea: we wish to make a prediction at x_0 given by $\hat{\mu}_0 = \hat{\alpha} + \hat{\beta} x_0$:

$$E[\hat{\mu}_0] = \alpha + \beta x_0 = \mu_0$$
$$Var(\hat{\mu}_0) = \sigma^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right),$$

and:

$$\frac{\hat{\mu}_0 - \mu_0}{\sqrt{S^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right)}} \sim t_{n-2}.$$

Brief review: Constructing confidence intervals

We can then use the pivotal quantities to construct $(1-\alpha) \times 100\%$ confidence intervals:

$$\frac{\hat{\mu}-\mu_0}{\sqrt{S^2 \cdot C}} \sim t_{n-2}$$

$$-t_{\alpha/2,n-2} < \frac{\hat{\mu}-\mu_0}{\sqrt{S^2 \cdot C}} < t_{\alpha/2,n-2}$$

$$-t_{\alpha/2,n-2}\sqrt{S^2 \cdot C} < \hat{\mu}-\mu_0 < t_{\alpha/2,n-2}\sqrt{S^2 \cdot C}$$

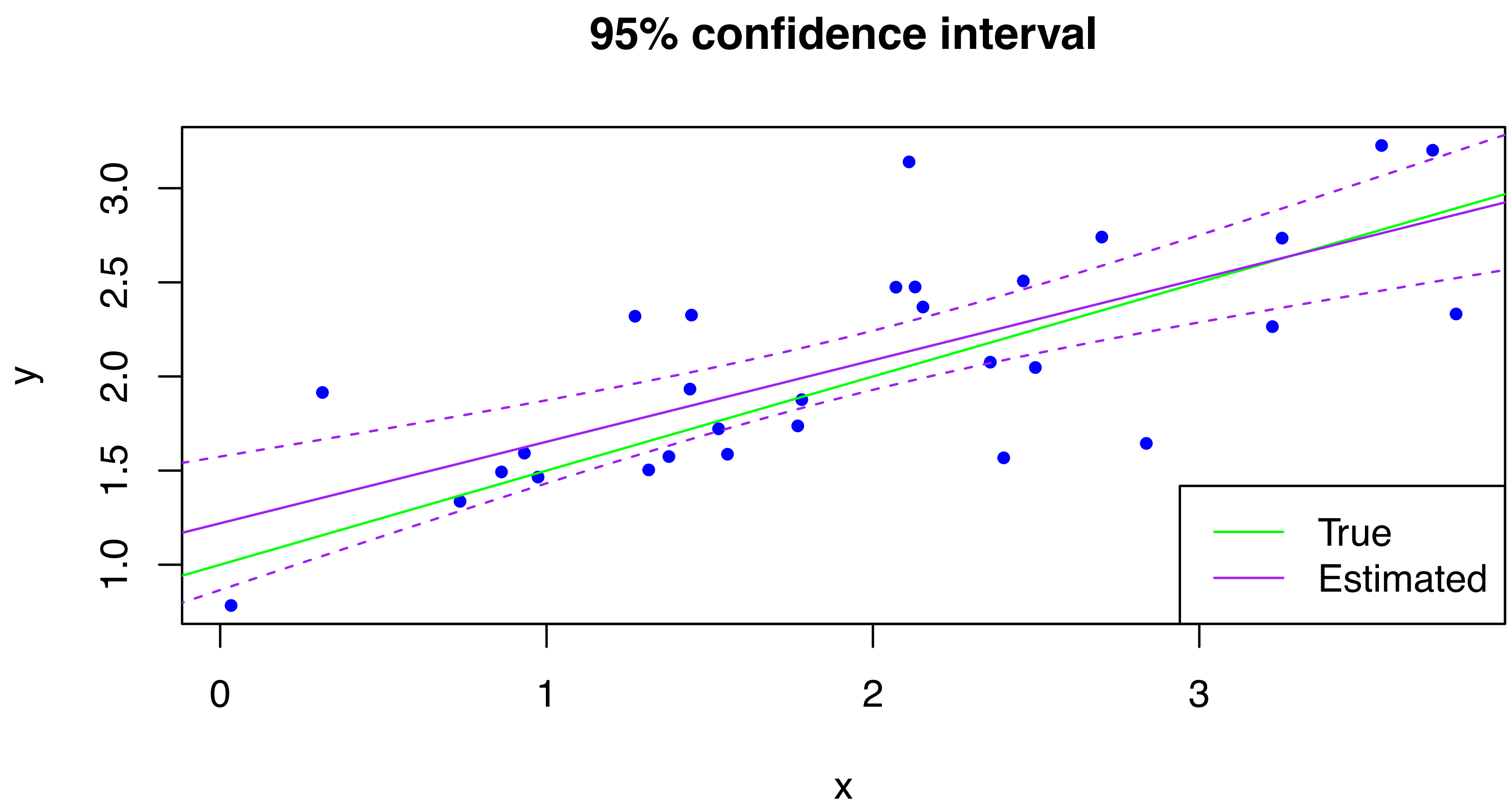
$$-\hat{\mu}-t_{\alpha/2,n-2}\sqrt{S^2 \cdot C} < -\mu_0 < -\hat{\mu}+t_{\alpha/2,n-2}\sqrt{S^2 \cdot C}$$

$$\hat{\mu}-t_{\alpha/2,n-2}\sqrt{S^2 \cdot C} < \mu_0 < \hat{\mu}+t_{\alpha/2,n-2}\sqrt{S^2 \cdot C},$$

for

$$C = \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}.$$

Brief review: Visualizing confidence intervals for μ_0

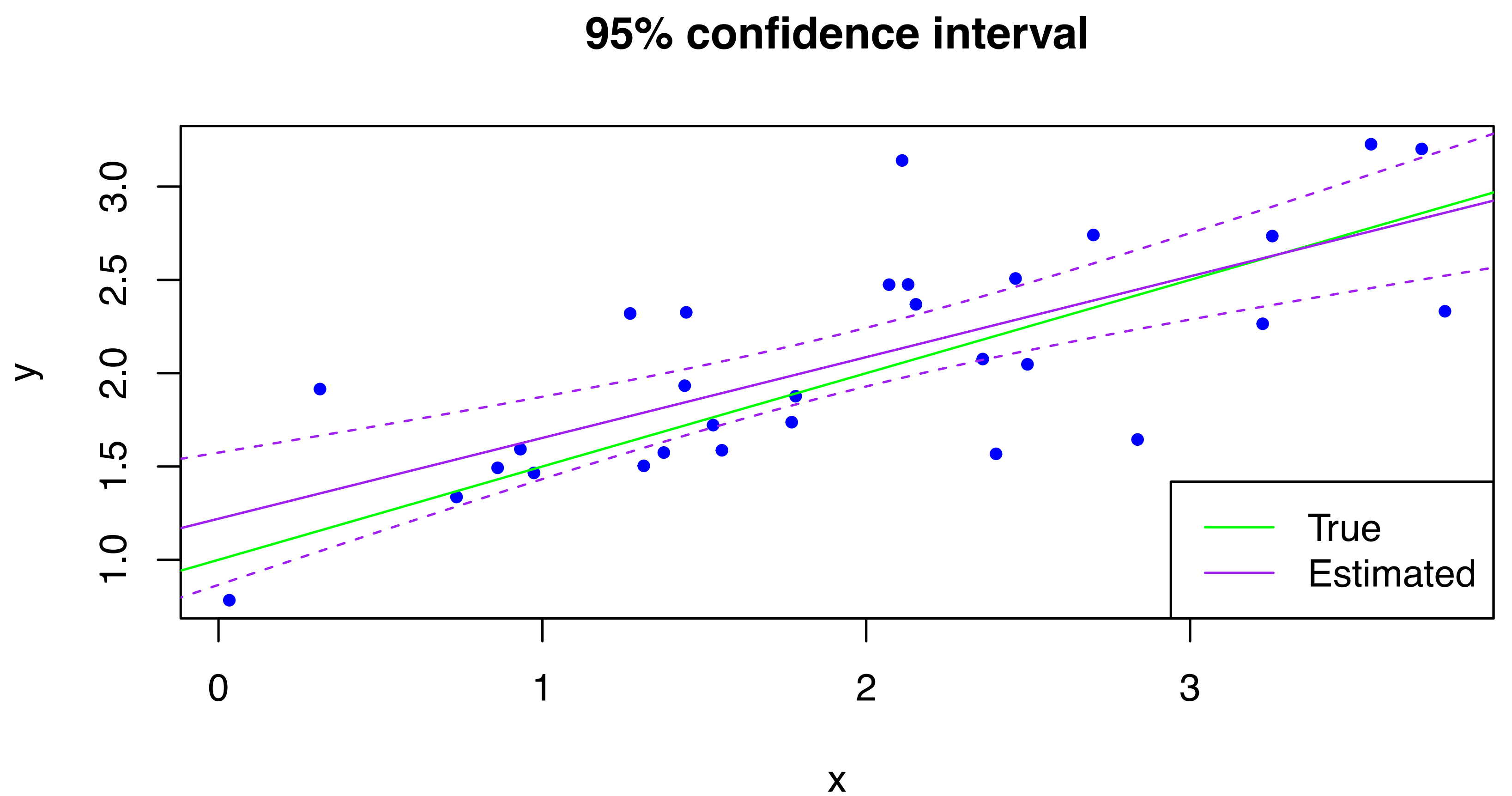


The confidence interval (CI) margin of error (MOE) is then:

$$t_{\alpha/2,n-2}\sqrt{S^2 \cdot C} = t_{\alpha/2,n-2}\sqrt{S^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}$$

.

Brief review: Visualizing confidence intervals for μ_0

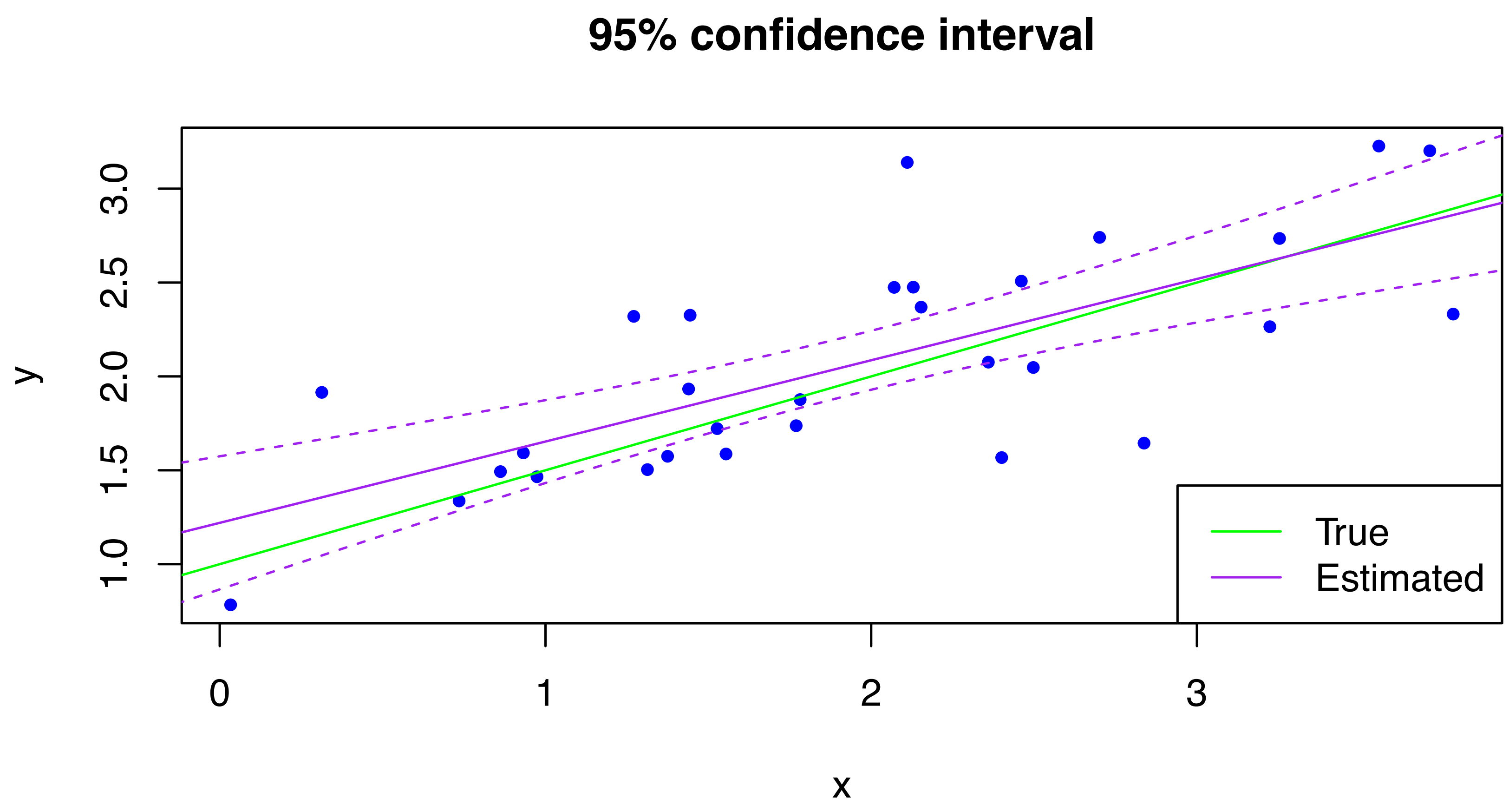


The confidence interval (CI) margin of error (MOE) is then:

$$t_{\alpha/2,n-2}\sqrt{S^2 \cdot C} = t_{\alpha/2,n-2}\sqrt{S^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}$$

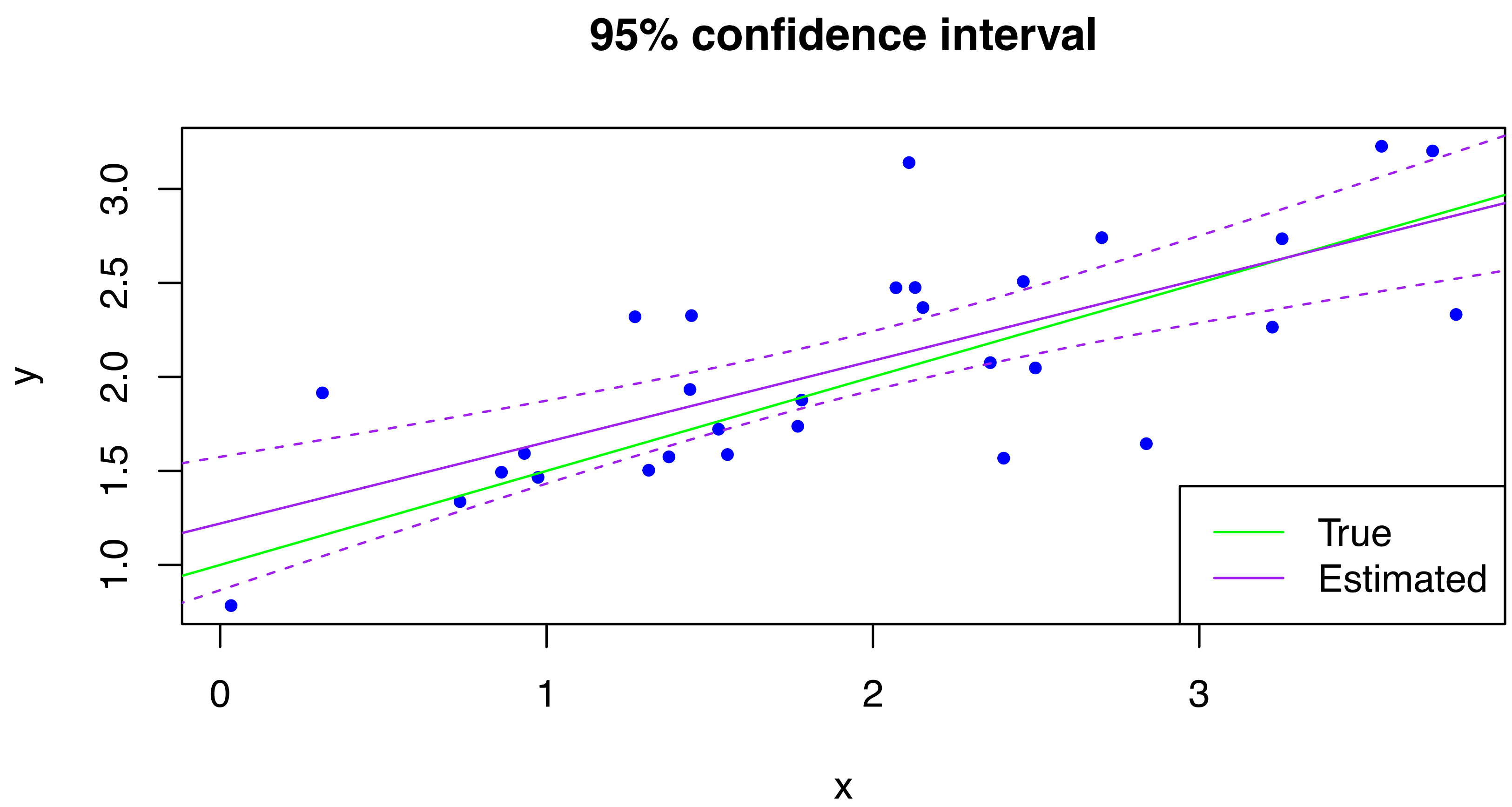
$$= t_{\alpha/2,n-2}S\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}.$$

Brief review: Predicting mean vs individual responses



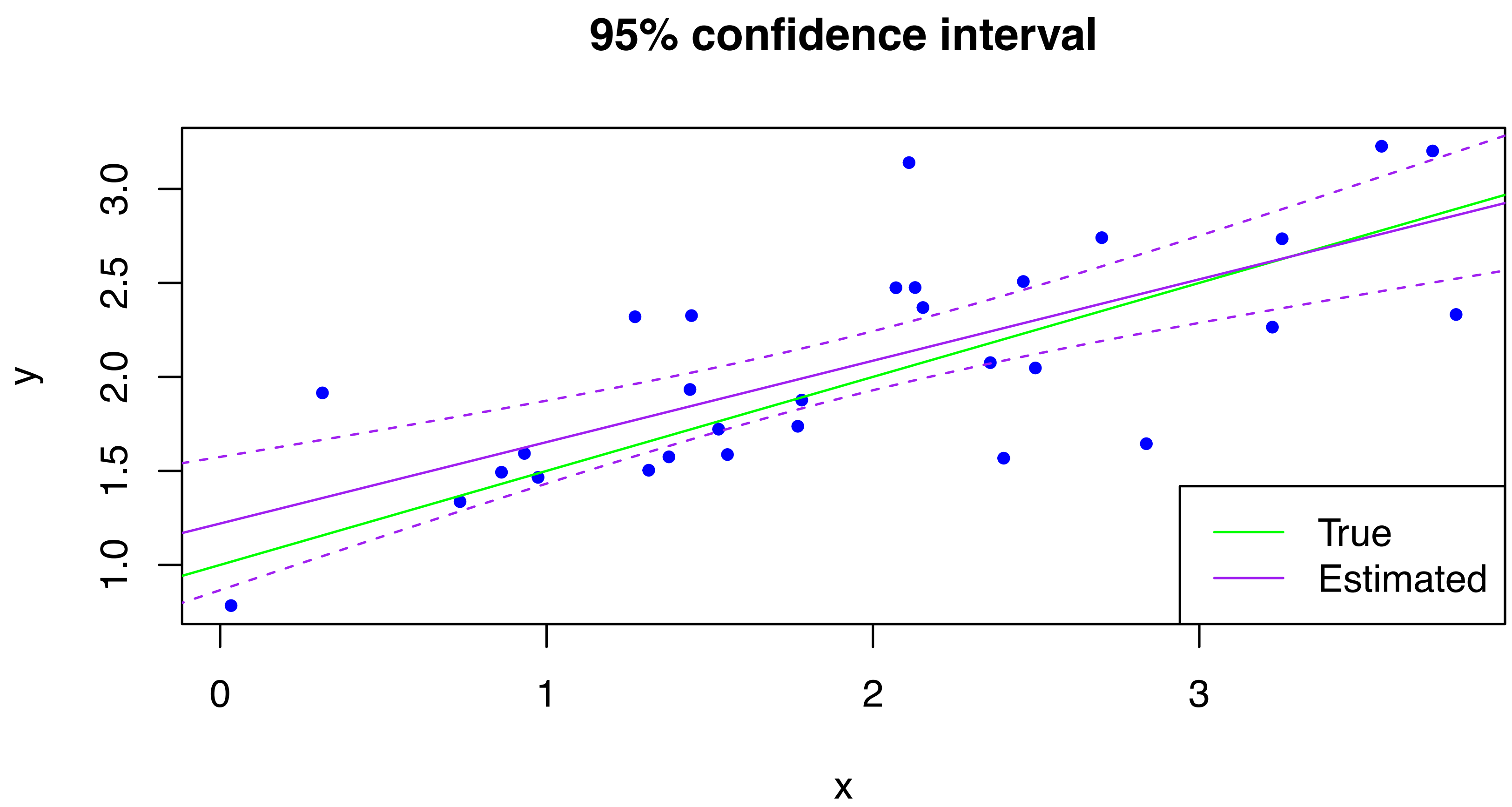
- Note that many of the individual responses are outside of the CI

Brief review: Predicting mean vs individual responses



- Note that many of the individual responses are outside of the CI
- That is because the CI is for the mean, μ_0 at some x_0, not a response Y_0 at some x_0!

Brief review: Predicting mean vs individual responses



- Note that many of the individual responses are outside of the CI
- That is because the CI is for the mean, μ_0 at some x_0, not a response Y_0 at some x_0!
- How can we make a CI for Y_0?

Brief review: Prediction interval setup

We will assume $\varepsilon_1, \ldots, \varepsilon_n \overset{iid}{\sim} N(0, \sigma^2)$, and,

$$Y_0 \sim N(\alpha + \beta x_0, \sigma^2),$$

and will estimate Y_0 with:

$$\hat{Y}_0 = \hat{\mu}_0 = \hat{\alpha} + \hat{\beta} x_0.$$

Hence,

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\right)$$

Brief review: Prediction interval setup

We will assume $\varepsilon_1, \ldots, \varepsilon_n \overset{iid}{\sim} N(0, \sigma^2)$, and,

$$Y_0 \sim N(\alpha + \beta x_0, \sigma^2),$$

and will estimate Y_0 with:

$$\hat{Y}_0 = \hat{\mu}_0 = \hat{\alpha} + \hat{\beta} x_0.$$

Hence,

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)\right)$$

- Note: this does not yet account for variability in responses around μ_0!

Brief review: Prediction intervals

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\right)$$

,

Brief review: Prediction intervals

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\right)$$

Note that $Var(Y_0 - \mu_0) = \sigma^2$.

,

Brief review: Prediction intervals

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\right)$$

Note that $Var(Y_0 - \mu_0) = \sigma^2$. Hence:

$$Y_0 - \hat{Y}_0 \sim N\left(\mu_0, \sigma^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)\right),$$

Brief review: Prediction intervals

$$\hat{Y}_0 = \hat{\mu}_0 \sim N\left(\mu_0, \sigma^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)\right)$$

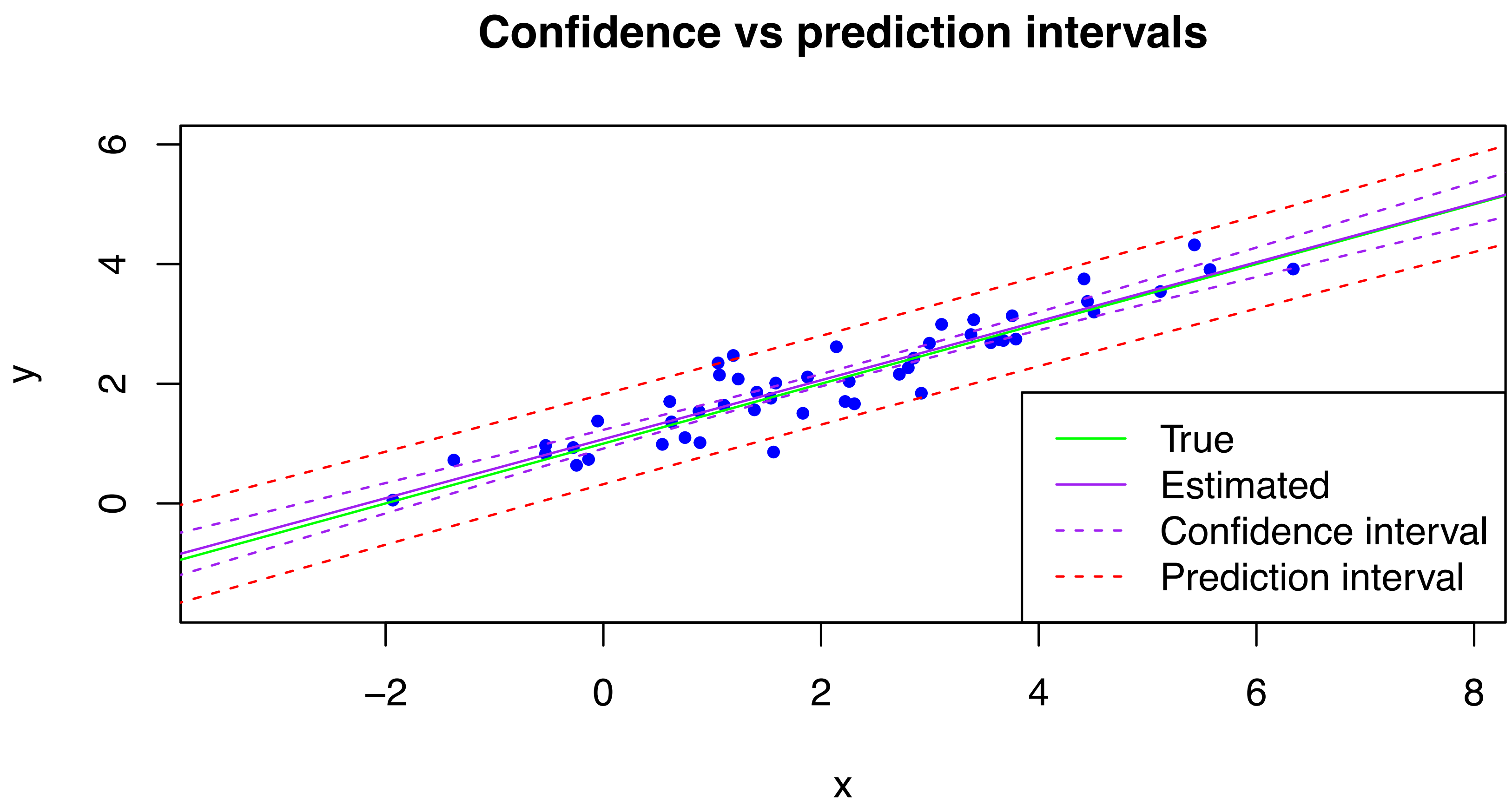
Note that $Var(Y_0 - \mu_0) = \sigma^2$. Hence:

$$Y_0 - \hat{Y}_0 \sim N\left(\mu_0, \sigma^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)\right),$$

yielding

$$T = \frac{Y_0 - \hat{Y}_0}{\sqrt{S^2\left(1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)}} \sim t_{n-2}.$$

Brief review: Predicting mean vs individual responses

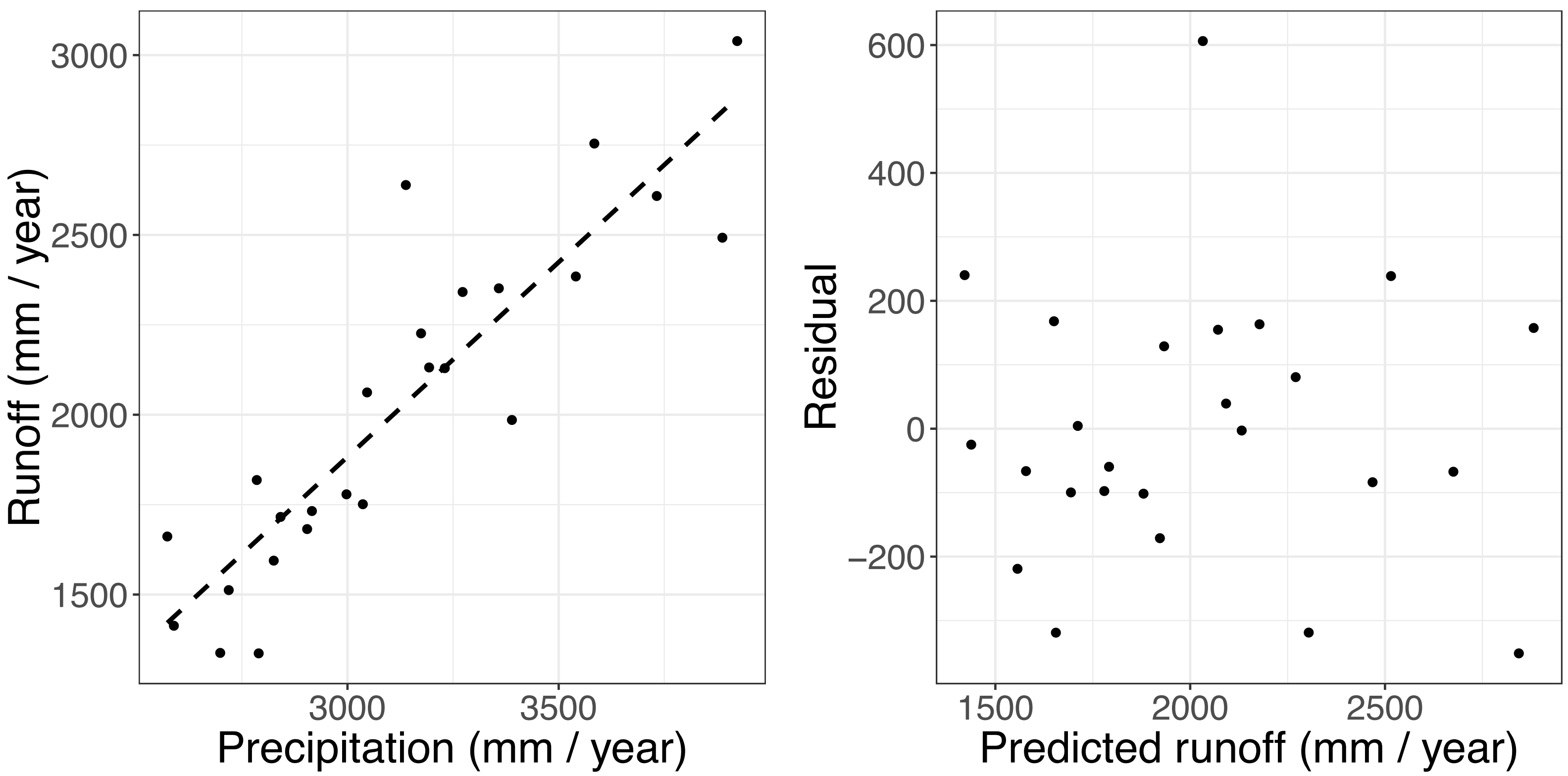


Margin of error (MOE) for CI vs prediction interval (PI):

$$\text{for CI:} \quad t_{\alpha/2,n-2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$\text{for PI:} \quad t_{\alpha/2,n-2} S \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

Runoff (based on problem 3, Spring 2019)

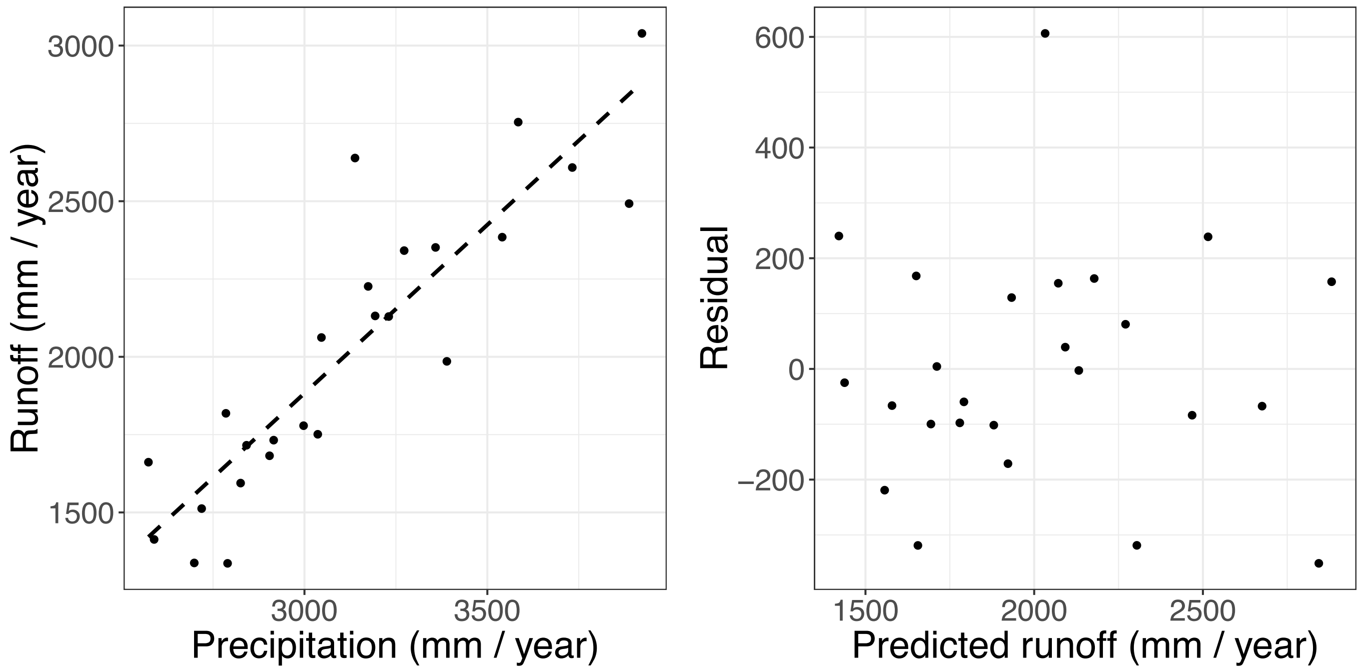


Assume the following linear relationship between annual runoff Y and annual precipitation x within a drainage basin,

$$Y = \beta_0 + \beta_1 x + \varepsilon, \tag{1}$$

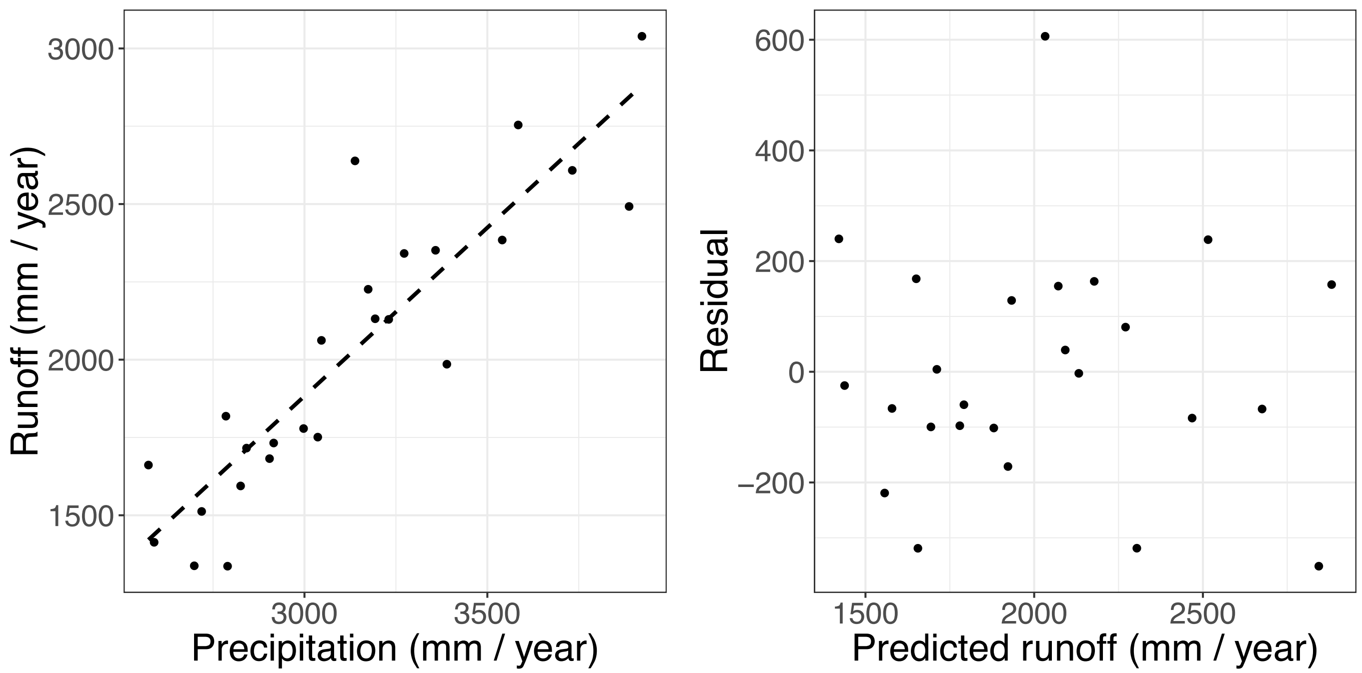
where β_0 and β_1 are unknown constants and ε is normally distributed with expected value (mean) 0 and unknown variance σ^2.

Runoff (problem 3b, Spring 2019)



Assume that we now are interested in predicting future runoff for a new year μ_0 given annual precipitation $x = x_0$, from the model defined in (1), where (x_0, Y_0) is independent of $(x_1, Y_1), (x_2, Y_2)$, $\ldots, (x_{25}, Y_{25})$. Assume that $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$ is a reasonable point estimator for the expected runoff $\mu_{Y|x_0} = \beta_0 + \beta_1 x_0$ when annual precipitation is x_0. Assume $\hat{\beta}_0 = -1364$, $\hat{\beta}_1 = 1.08$, and $S^2 = 156^2$.

Runoff (problem 3b, Spring 2019)



Assume $\hat{\beta}_0 = -1364$, $\hat{\beta}_1 = 1.08$, and $S^2 = 156^2$. Questions:

1. Calculate the predicted runoff for a year with annual precipitation equal to $x_0 = 2000$ mm/yr.
2. Show that $E[\hat{Y}_0 - Y_0] = 0$.
3. Show that $Var(\hat{Y}_0 - Y_0) = \sigma^2 \left(1 + \frac{1}{25} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{25}(x_i - \bar{x})^2}\right)$.
4. Give a 95% prediction interval for the observed annual runoff for a year with precipitation equal to x_0.

Medicine concentration (based on problem 3, Spring 2016)

In this problem we are going to consider a disease for which the treatment is to inject a medicine into the blood. Let x denote the dose of medicine a patient gets injected. We will assume that this dose can be controlled and therefore we do not consider x to be stochastic. Twenty-four hours after the medicine is injected, the concentration, Y, of the medicine in the blood is measured. We assume the following linear regression model for the relation between x and Y,

$$Y = bx + \varepsilon,$$

where b is an unknown parameter and ε has a normal distribution with zero mean and variance σ^2. In all parts of this problem we assume $\sigma^2 = 2.0^2$ to be known.

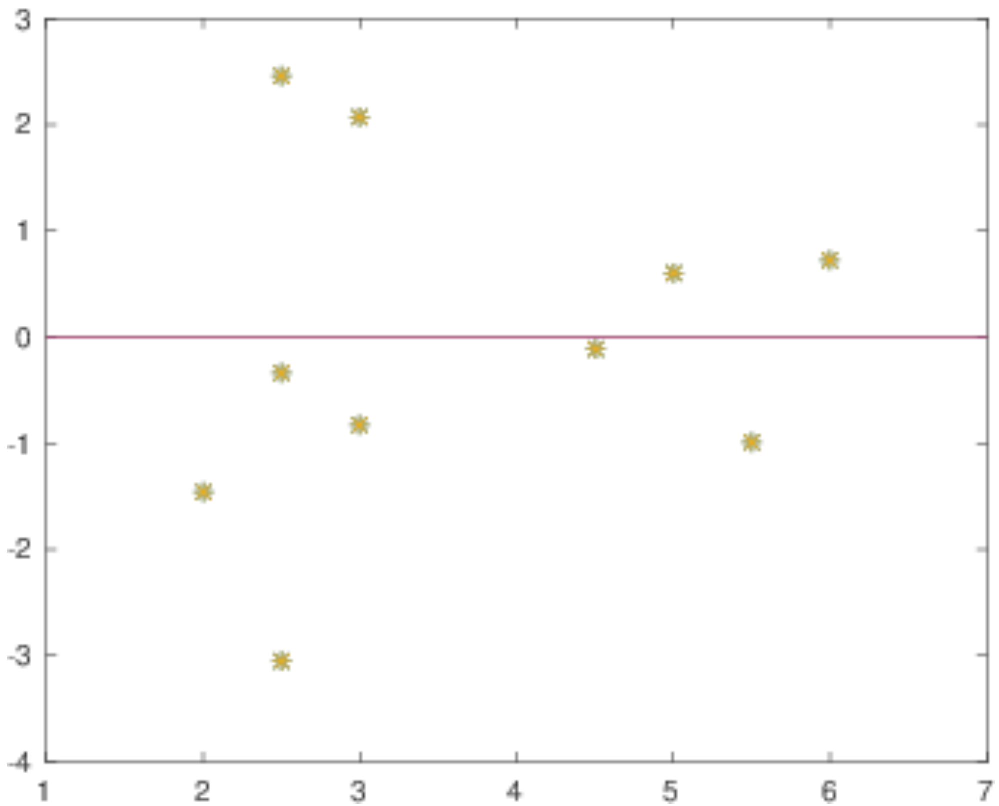
Medicine concentration (based on problem 3, Spring 2016)

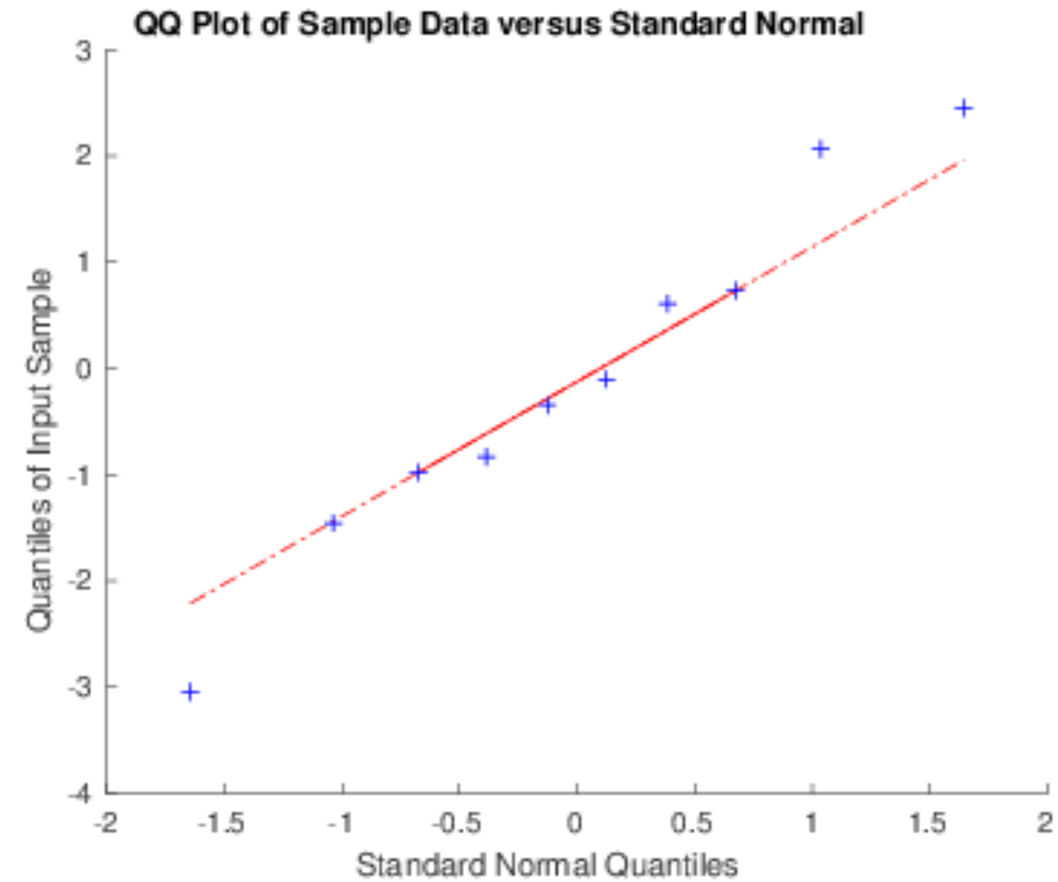
Assume we have observed values for $n = 10$ patients and let x_i and Y_i for $i = 1, 2, \ldots, n$ be corresponding values of injected dose and measured concentration in the blood. We assume $Y_1, Y_2, \ldots, Y_n$ to be independent stochastic variables and we assume the regression model above to hold for each of them. Observed values for the n patients are:

patient i	1	2	3	4	5	6	7	8	9	10
x_i	2.0	5.0	2.5	2.5	4.5	5.5	3.0	6.0	3.0	2.5
y_i	2.9	11.5	5.1	7.9	9.7	11.0	8.6	13.8	5.7	2.4

It is given that $\sum_{i=1}^{n} x_i = 36.5$, $\sum_{i=1}^{n} x_i^2 = 152.25$ and $\sum_{i=1}^{n} x_i y_i = 331.65$.

Medicine concentration (based on problem 3, Spring 2016)



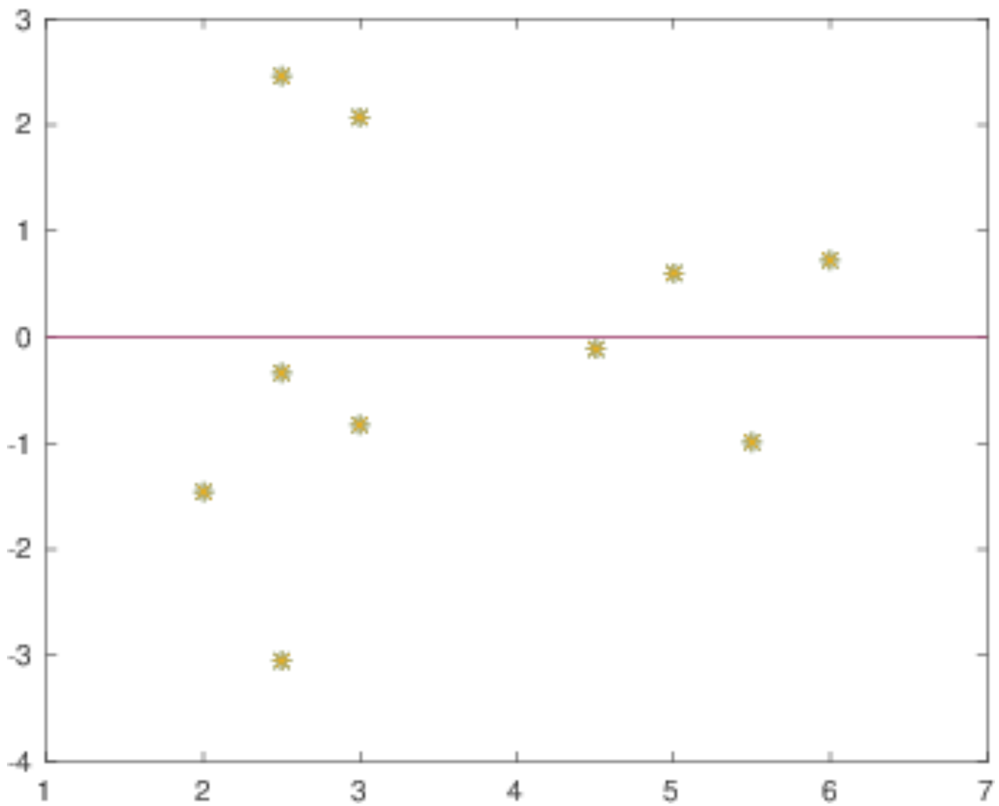


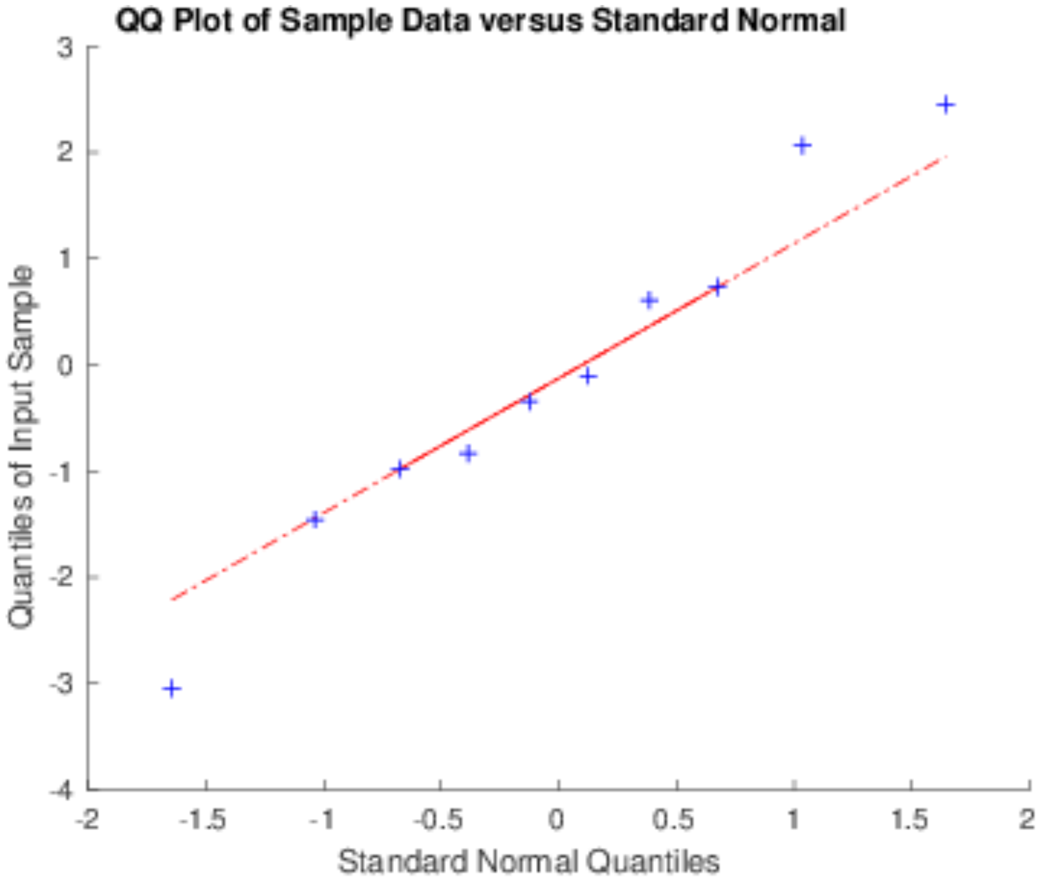
Answer as an individual:

1. How should the residual plot (left) and the normal quantile-quantile (QQ) plot (right), be interpreted?

01:00

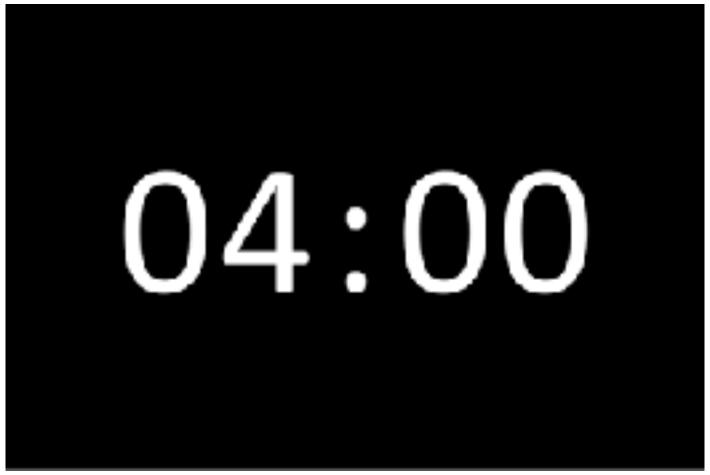
Medicine concentration (based on problem 3, Spring 2016)



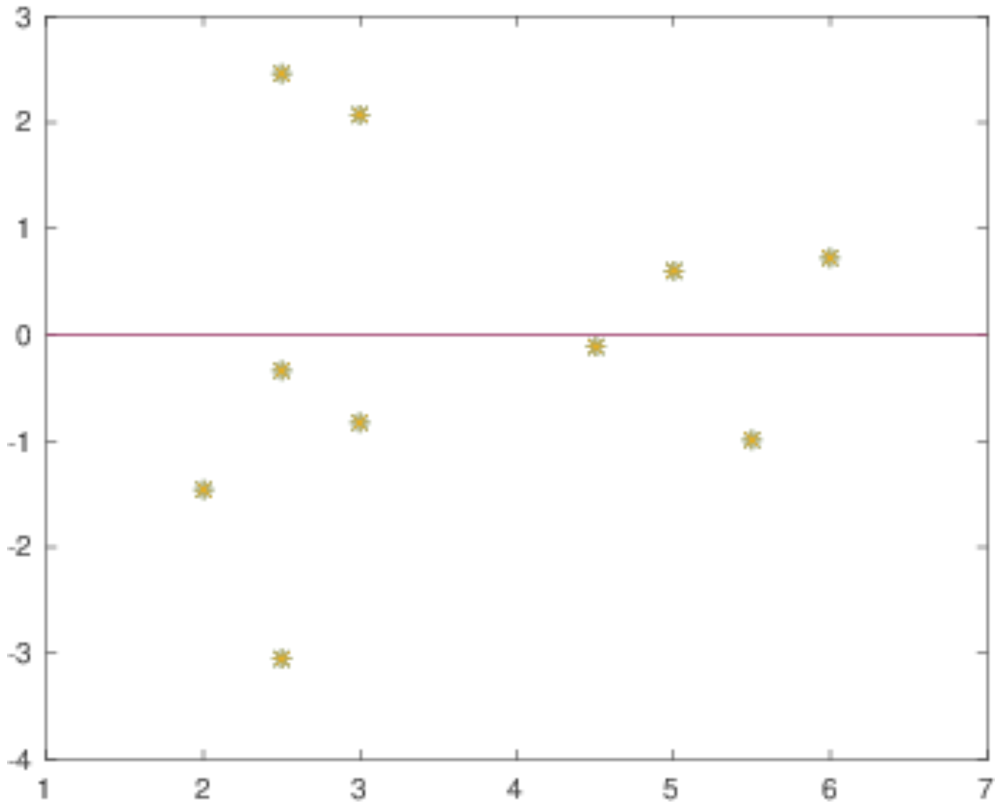


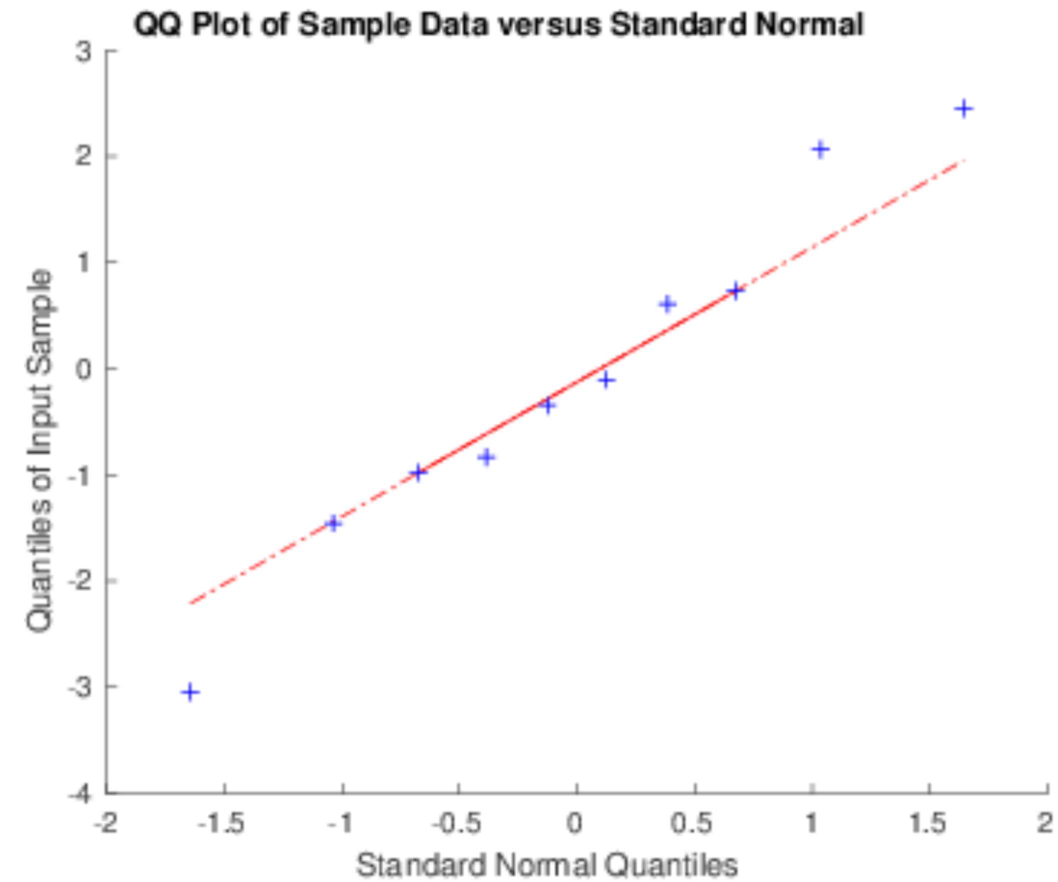
Discuss in a group:

1. How should the residual plot (left) and the normal quantile-quantile (QQ) plot (right), be interpreted?



Medicine concentration (based on problem 3, Spring 2016)



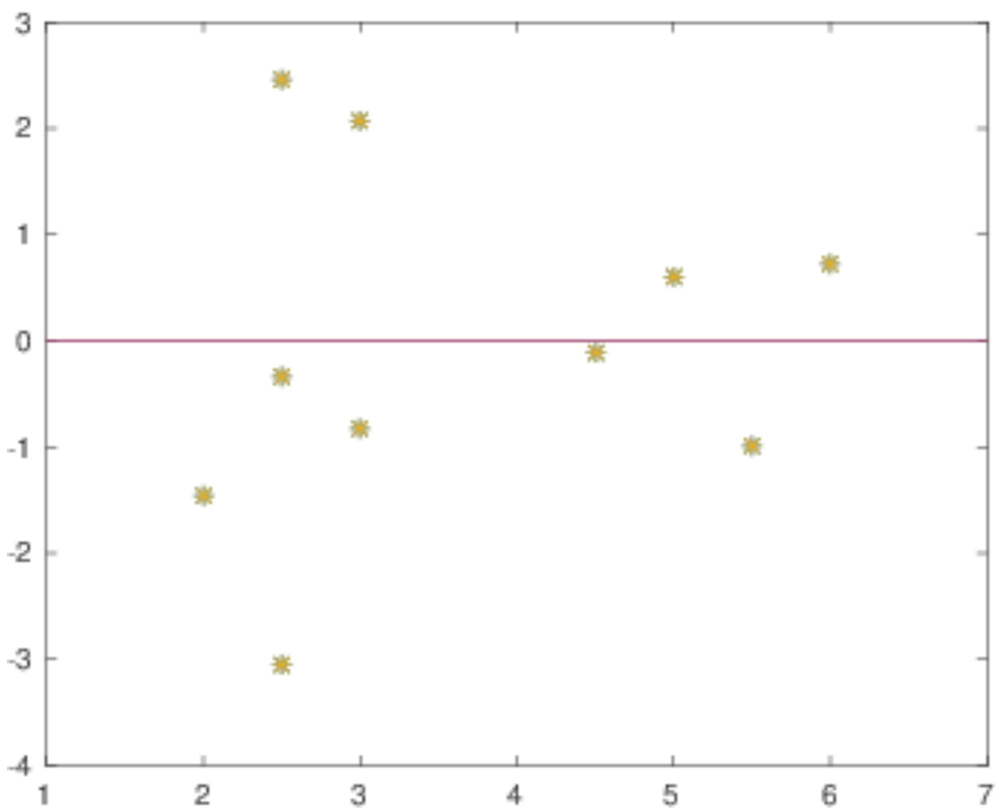


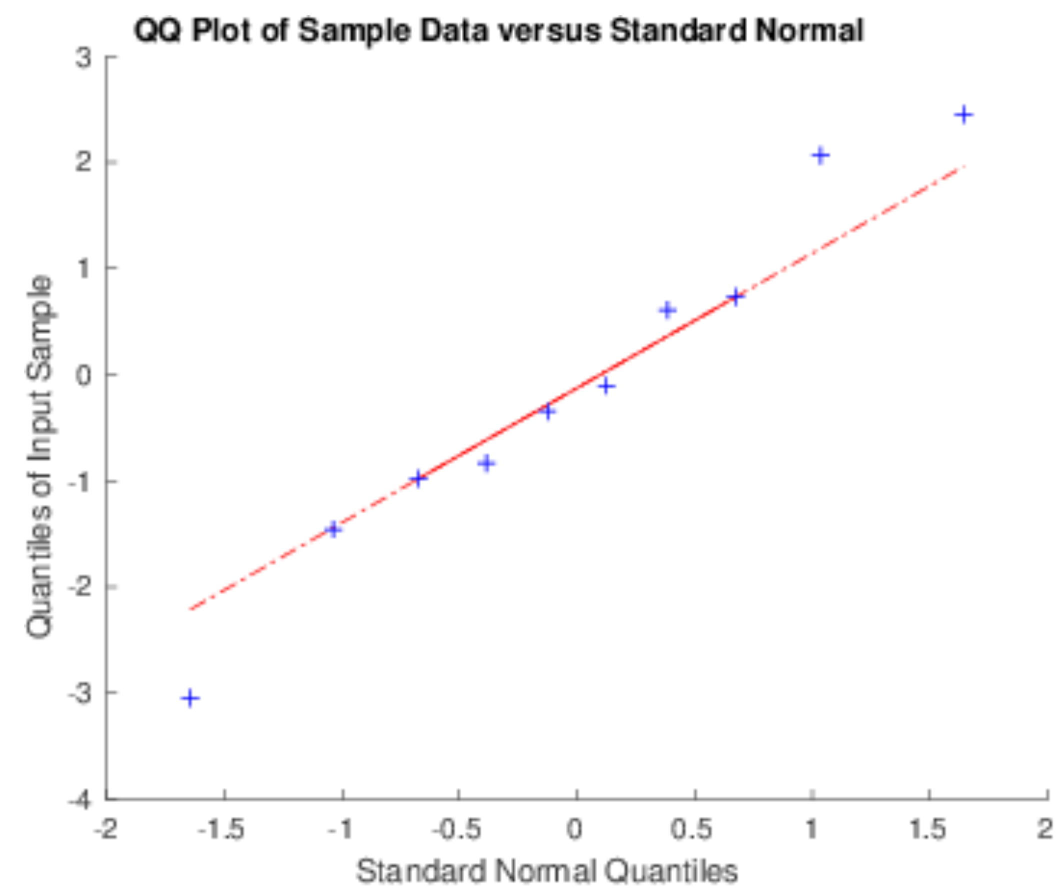
Answer as an individual:

2. Given the residual plot (left) and the normal quantile-quantile (QQ) plot for the residuals (right), do the model assumptions appear to be reasonable?

01:00

Medicine concentration (based on problem 3, Spring 2016)





Discuss in a group:

2. Given the residual plot (left) and the normal quantile-quantile (QQ) plot for the residuals (right), do the model assumptions appear to be reasonable?

03:00

Medicine concentration (based on problem 3, Spring 2016)

$$Y = bx + \varepsilon$$

To estimate the parameter b, the estimators

$$\widetilde{b} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} x_i} \quad , \quad \widehat{b} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \text{and} \quad \widehat{\widehat{b}} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$$

are proposed. It is given that $\widehat{\widehat{b}}$ is unbiased and that $\mathrm{Var}\left[\widehat{\widehat{b}}\right] = \sigma^2 / \sum_{i=1}^{n} x_i^2$.

It is given that $\sum_{i=1}^{n} x_i = 36.5$, $\sum_{i=1}^{n} x_i^2 = 152.25$ and $\sum_{i=1}^{n} x_i y_i = 331.65$.

Discuss and solve as a group:

3. Which of the three estimators do you prefer when $n = 10$?

A) $\tilde{b}$

B) $\hat{b}$

C) $\hat{\hat{b}}$

07:00

Medicine concentration (based on problem 3, Spring 2016)

$$Y = bx + \varepsilon$$

To estimate the parameter b, the estimators

$$\widetilde{b} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i} \quad , \quad \widehat{b} = \frac{1}{n}\sum_{i=1}^n Y_i \quad \text{and} \quad \widehat{\widehat{b}} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$$

are proposed. It is given that $\widehat{\widehat{b}}$ is unbiased and that $\text{Var}\left[\widehat{\widehat{b}}\right] = \sigma^2 / \sum_{i=1}^n x_i^2$.

It is given that $\sum_{i=1}^n x_i = 36.5$, $\sum_{i=1}^n x_i^2 = 152.25$ and $\sum_{i=1}^n x_i y_i = 331.65$.

Discuss and solve as a group:

3. Which of the three estimators do you prefer when $n = 10$?

A) $\tilde{b}$

B) $\hat{b}$

C) $\hat{\hat{b}}$

Medicine concentration (based on problem 3, Spring 2016)

$$Y = bx + \varepsilon$$

$$\hat{\hat{b}} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$$

$$\phi(y;\ \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2}(y-\mu)^2 \right\}$$

Discuss and solve as a group:

3. Write down the likelihood function for b given observation pairs $(x_1, y_1), \ldots, (x_n, y_n)$ and given $\sigma^2 = 2^2$. What is the maximum likelihood estimator for b?

A) $\hat{b}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} Y_i$

B) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} x_i}$

C) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$

D) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) Y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$

08:00

Medicine concentration (based on problem 3, Spring 2016)

$$Y = bx + \varepsilon$$

$$\hat{\hat{b}} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$$

$$\phi(y;\ \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}$$

Discuss and solve as a group:

3. Write down the likelihood function for b given observation pairs $(x_1, y_1), \ldots, (x_n, y_n)$ and given $\sigma^2 = 2^2$. What is the maximum likelihood estimator for b?

A) $\hat{b}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} Y_i$

B) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} x_i}$

C) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$

D) $\hat{b}_{\text{MLE}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Medicine concentration (based on problem 3, Spring 2016)

It is given that $\sum_{i=1}^{n} x_i = 36.5$, $\sum_{i=1}^{n} x_i^2 = 152.25$ and $\sum_{i=1}^{n} x_i y_i = 331.65$.

Discuss and solve as a group:

4. Derive a $(1-\alpha) \times 100\%$ confidence interval for b as a function of $x_1, \ldots, x_n$, $Y_1, \ldots, Y_n$, n, σ^2, and α, assuming $\sigma^2 = 2^2$, $\alpha = 0.1$, and $n = 10$.

A) $(1.9117, 2.4450)$

B) $(1.8769, 2.4797)$

C) $(1.8812, 2.4755)$

D) $(1.8845, 2.4721)$

06:00

Medicine concentration (based on problem 3, Spring 2016)

It is given that $\sum_{i=1}^{n} x_i = 36.5$, $\sum_{i=1}^{n} x_i^2 = 152.25$ and $\sum_{i=1}^{n} x_i y_i = 331.65$.

Discuss and solve as a group:

4. Derive a $(1-\alpha) \times 100\%$ confidence interval for b as a function of $x_1, \ldots, x_n, Y_1, \ldots, Y_n, n, \sigma^2$, and α, assuming $\sigma^2 = 2^2$, $\alpha = 0.1$, and $n = 10$.

A) $(1.9117, 2.4450)$

B) $(1.8769, 2.4797)$

C) $(1.8812, 2.4755)$

D) $(1.8845, 2.4721)$

Medicine concentration (based on problem 3, Spring 2016)

It is important that the concentration of medicine in the blood is not too high, as this may give serious adverse effects. After having observed the results for the first $n = 10$ patients (given above), the medical doctors receive a new patient and after having examined this patient the medical doctors decide that it is important that the measured concentration of medicine in the blood of this patient does not exceed 10.0.

e) Find the highest dose x_0 this new patient can get injected if one requires a probability of at least 0.95 for the event that the measured concentration of medicine after twenty-four hours does not exceed 10.0.

Solve yourself!

Reminders

- Summary sessions on Wednesday and Monday
- Stack and homework due Friday 11:59 pm