THE TASK

You are assigned a simulation design and a dataset from the Wooldridge package based on the last digit of your student ID number. Submitting a report for a wrong dataset results in the mark of zero for this assignment. This computer assignment corresponds to 10% of your final grade. The details about individual dataset and models are provided on the pages below.

Your answers have to be reported in the following order.

STUDENT ID NUMBER: (*insert here*)

1. Plot for Question 1
2. Plot with two histograms for Question 2
3. Estimated marginal effects with interpretation.

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		

You have to submit a single pdf file. For all the numbers report the first 3 digits after the decimal point: 3.567 instead of 3.6. You then have to include your R or MATLAB code. *Hint: in RStudio go to* `File - Knit Document` *and include the generated report. The code will compile if and only if it is written without mistakes. If your code does not compile you still have to include it (just copy and paste as a text after the table).*

The submission deadline for this assignment is **12.00hrs Greenwich Mean Time January 20, 2021**. The report has to be submitted via Turnitin on Blackboard. The assignment will be marked in accordance with the general *SoSS PG Marking Criteria* (available on Blackboard). Please make sure you are familiar with the University's rules and regulations regarding plagiarism.

In general your task is to produce an empirical analysis of binary response models.

STUDENT ID NUMBERS ENDING WITH 0 OR 1

- ID Example: 123450 or 123451.
- Use the Wooldridge dataset `affairs` to estimate the **logit** model
- Use `help(affairs)` to get the variable description.

1. Create a figure with two plots alongside each other: Plot 1: the log likelihood function for a Bernoulli distributed random variable together with Plot 2: the numerical derivative (the gradient) of the log likelihood function. Proceed as follows:
   - Simulate a vector x of length $n = 100$ from a Bernoulli distribution using `rbinom(n,1,`θ_0`)` with $\theta_0 = 0.2$.
   - Write a function which takes x and θ as an input and returns the value of the log likelihood function $LLF_n(\theta)$ as an output. Evaluate this function on the grid of $\theta = \{0.01, 0.02, ..., 0.97, 0.98, 0.99\}$. Use the Example 6.1 from the lecture notes a guidance.
   - Create a vector $dx = \theta_i - \theta_{i-1}$ for $i = 2, ..., 99$. Create a vector of the corresponding changes in the log likelihood function $dy = LLF_n(\theta_i) - LLF_n(\theta_{i-1})$. Define the numerical derivative as dy/dx. Use the homework exercise from the computer tutorial 1 as guidance.

   - Create a figure with two subplots next to each other. On the left-hand side panel plot the log likelihood function using the grid of $\theta = \{0.01, 0.02, ..., 0.98, 0.99\}$ as x-axis. Add a vertical line at θ_0. On the left-hand side plot the numerical derivative of the log-likelihood function with the grid of θ as x-axis. Add a vertical line at θ_0 and a horizontal line at zero.

2. Create a figure with two histograms alongside each other: Plot 1: a histogram of the centered maximum likelihood estimator (MLE) of the sample mean of a normally distributed variable and Plot 2: a histogram of the standardized MLE of the sample mean of a normally distributed variable. Proceed as follows:

   - Generate $MC = 5000$ simulation draws of $x_i \sim \mathcal{N}(\mu_0, \sigma_0)$, $i = 1, ..., 100$. Use $\mu_0 = 3$, $\sigma_0 = 100$.
   - For every simulation draw, compute the MLE of the mean $\hat{\mu}_{ML}$. Use Example 6.2 from the lecture notes as guidance.
   - Using the results of Example 6.2 compute the standardization of $\hat{\mu}_{ML}$, which has a limiting standard normal distribution $\mathcal{N}(0, 1)$.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot a histogram of the centered MLE $\hat{\mu}_{ML} - \mu_0$. On the right-hand side panel plot a histogram of the standardized MLE from the previous step. Plot the theoretical density of the standard normal distribution $\mathcal{N}(0, 1)$ on top of each histogram.

3. Estimate a logistic regression with the dependent variable $y = affair$ and explanatory variables $x_1 = ratemarr$, $x_2 = kids$, $x_3 = relig$. Using the library `mfx` calculate the estimated marginal effect for every explanatory variable using the option `atmean = TRUE`. Write down the economic interpretation of the estimated marginal effects, which are significantly different from zero at 5% significance level. Use computer tutorial 8 together with Section 6.2.4 from the lecture notes as guidance. Report your answers in the following table:

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		

STUDENT ID NUMBERS ENDING WITH 2 OR 3

- ID Example: 123452 or 123453.
- Use the Wooldridge dataset `alcohol` to estimate the **probit** model
- Use `help(alcohol)` to get the variable description.

1. Create a figure with two plots alongside each other: Plot 1: the log likelihood function for a Bernoulli distributed random variable together with Plot 2: the numerical derivative (the gradient) of the log likelihood function. Proceed as follows:
   - Simulate a vector x of length $n = 100$ from a Bernoulli distribution using `rbinom(n,1,`θ_0`)` with $\theta_0 = 0.4$.
   - Write a function which takes x and θ as an input and returns the value of the log likelihood function $LLF_n(\theta)$ as an output. Evaluate this function on the grid of $\theta = \{0.01, 0.02, ..., 0.97, 0.98, 0.99\}$. Use the Example 6.1 from the lecture notes a guidance.
   - Create a vector $dx = \theta_i - \theta_{i-1}$ for $i = 2, ..., 99$. Create a vector of the corresponding changes in the log likelihood function $dy = LLF_n(\theta_i) - LLF_n(\theta_{i-1})$. Define the numerical derivative as dy/dx. Use the homework exercise from the computer tutorial 1 as guidance.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot the log likelihood function using the grid of $\theta = \{0.01, 0.02, ..., 0.98, 0.99\}$ as x-axis. Add a vertical line at θ_0. On the left-hand side plot the numerical derivative of the log-likelihood function with the grid of θ as x-axis. Add a vertical line at θ_0 and a horizontal line at zero.

2. Create a figure with two histograms alongside each other: Plot 1: a histogram of the centered maximum likelihood estimator (MLE) of the sample mean of a normally distributed variable and Plot 2: a histogram of the standardized MLE of the sample mean of a normally distributed variable. Proceed as follows:
   - Generate $MC = 5000$ simulation draws of $x_i \sim \mathcal{N}(\mu_0, \sigma_0)$, $i = 1, ..., 100$. Use $\mu_0 = 1$, $\sigma_0 = 3$.
   - For every simulation draw, compute the MLE of the mean $\hat{\mu}_{ML}$. Use Example 6.2 from the lecture notes as guidance.
   - Using the results of Example 6.2 compute the standardization of $\hat{\mu}_{ML}$, which has a limiting standard normal distribution $\mathcal{N}(0, 1)$.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot a histogram of the centered MLE $\hat{\mu}_{ML} - \mu_0$. On the right-hand side panel plot a histogram of the standardized MLE from the previous step. Plot the theoretical density of the standard normal distribution $\mathcal{N}(0, 1)$ on top of each histogram.

3. Estimate a probit model with the dependent variable $y = employ$ and explanatory variables $x_1 = age$, $x_2 = educ$, $x_3 = abuse$. Using the library `mfx` calculate the estimated marginal effect for every explanatory variable using the option `atmean = TRUE`. Write down the economic interpretation of the estimated marginal effects, which are significantly different from zero at 5% significance level. Use computer tutorial 8 together with Section 6.2.4 from the lecture notes as guidance. Report your answers in the following table:

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		

STUDENT ID NUMBERS ENDING WITH 4 OR 5

- ID Example: 123454 or 123455.
- Use the Wooldridge dataset `mroz` to estimate the **logit** model
- Use `help(mroz)` to get the variable description.

1. Create a figure with two plots alongside each other: Plot 1: the log likelihood function for a Bernoulli distributed random variable together with Plot 2: the numerical derivative (the gradient) of the log likelihood function. Proceed as follows:
   - Simulate a vector x of length $n = 100$ from a Bernoulli distribution using `rbinom(n,1,`θ_0`)` with $\theta_0 = 0.6$.
   - Write a function which takes x and θ as an input and returns the value of the log likelihood function $LLF_n(\theta)$ as an output. Evaluate this function on the grid of $\theta = \{0.01, 0.02, ..., 0.97, 0.98, 0.99\}$. Use the Example 6.1 from the lecture notes a guidance.
   - Create a vector $dx = \theta_i - \theta_{i-1}$ for $i = 2, ..., 99$. Create a vector of the corresponding changes in the log likelihood function $dy = LLF_n(\theta_i) - LLF_n(\theta_{i-1})$. Define the numerical derivative as dy/dx. Use the homework exercise from the computer tutorial 1 as guidance.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot the log likelihood function using the grid of $\theta = \{0.01, 0.02, ..., 0.98, 0.99\}$ as x-axis. Add a vertical line at θ_0. On the left-hand side plot the numerical derivative of the log-likelihood function with the grid of θ as x-axis. Add a vertical line at θ_0 and a horizontal line at zero.

2. Create a figure with two histograms alongside each other: Plot 1: a histogram of the centered maximum likelihood estimator (MLE) of the sample mean of a normally distributed variable and Plot 2: a histogram of the standardized MLE of the sample mean of a normally distributed variable. Proceed as follows:

   - Generate $MC = 5000$ simulation draws of $x_i \sim \mathcal{N}(\mu_0, \sigma_0), i = 1, ..., 100$. Use $\mu_0 = -8, \sigma_0 = 0.1$.
   - For every simulation draw, compute the MLE of the mean $\hat{\mu}_{ML}$. Use Example 6.2 from the lecture notes as guidance.
   - Using the results of Example 6.2 compute the standardization of $\hat{\mu}_{ML}$, which has a limiting standard normal distribution $\mathcal{N}(0, 1)$.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot a histogram of the centered MLE $\hat{\mu}_{ML} - \mu_0$. On the right-hand side panel plot a histogram of the standardized MLE from the previous step. Plot the theoretical density of the standard normal distribution $\mathcal{N}(0, 1)$ on top of each histogram.

3. Estimate a logistic regression with the dependent variable $y = inlf$ and explanatory variables $x_1 = kidslt6, x_2 = educ, x_3 = exper$. Using the library `mfx` calculate the estimated marginal effect for every explanatory variable using the option `atmean = TRUE`. Write down the economic interpretation of the estimated marginal effects, which are significantly different from zero at 5% significance level. Use computer tutorial 8 together with Section 6.2.4 from the lecture notes as guidance. Report your answers in the following table:

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		

STUDENT ID NUMBERS ENDING WITH 6 OR 7

- ID Example: 123456 or 123457.
- Use the Wooldridge dataset `loanapp` to estimate the **probit** model
- Use `help(loanapp)` to get the variable description.

1. Create a figure with two plots alongside each other: Plot 1: the log likelihood function for a Bernoulli distributed random variable together with Plot 2: the numerical derivative (the gradient) of the log likelihood function. Proceed as follows:
   - Simulate a vector x of length $n = 100$ from a Bernoulli distribution using `rbinom(n,1,`θ_0`)` with $\theta_0 = 0.8$.
   - Write a function which takes x and θ as an input and returns the value of the log likelihood function $LLF_n(\theta)$ as an output. Evaluate this function on the grid of $\theta = \{0.01, 0.02, ..., 0.97, 0.98, 0.99\}$. Use the Example 6.1 from the lecture notes a guidance.
   - Create a vector $dx = \theta_i - \theta_{i-1}$ for $i = 2, ..., 99$. Create a vector of the corresponding changes in the log likelihood function $dy = LLF_n(\theta_i) - LLF_n(\theta_{i-1})$. Define the numerical derivative as dy/dx. Use the homework exercise from the computer tutorial 1 as guidance.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot the log likelihood function using the grid of $\theta = \{0.01, 0.02, ..., 0.98, 0.99\}$ as x-axis. Add a vertical line at θ_0. On the left-hand side plot the numerical derivative of the log-likelihood function with the grid of θ as x-axis. Add a vertical line at θ_0 and a horizontal line at zero.

2. Create a figure with two histograms alongside each other: Plot 1: a histogram of the centered maximum likelihood estimator (MLE) of the sample mean of a normally distributed variable and Plot 2: a histogram of the standardized MLE of the sample mean of a normally distributed variable. Proceed as follows:
   - Generate $MC = 5000$ simulation draws of $x_i \sim \mathcal{N}(\mu_0, \sigma_0)$, $i = 1, ..., 100$. Use $\mu_0 = -42$, $\sigma_0 = 42$.
   - For every simulation draw, compute the MLE of the mean $\hat{\mu}_{ML}$. Use Example 6.2 from the lecture notes as guidance.
   - Using the results of Example 6.2 compute the standardization of $\hat{\mu}_{ML}$, which has a limiting standard normal distribution $\mathcal{N}(0, 1)$.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot a histogram of the centered MLE $\hat{\mu}_{ML} - \mu_0$. On the right-hand side panel plot a histogram of the standardized MLE from the previous step. Plot the theoretical density of the standard normal distribution $\mathcal{N}(0, 1)$ on top of each histogram.

3. Estimate a probit model with the dependent variable $y = approve$ and explanatory variables $x_1 = hrat$, $x_2 = self$, $x_3 = male$. Using the library `mfx` calculate the estimated marginal effect for every explanatory variable using the option `atmean = TRUE`. Write down the economic interpretation of the estimated marginal effects, which are significantly different from zero at 5% significance level. Use computer tutorial 8 together with Section 6.2.4 from the lecture notes as guidance. Report your answers in the following table:

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		

STUDENT ID NUMBERS ENDING WITH 8 OR 9

- ID Example: 123458 or 123459.
- Use the Wooldridge dataset `crime1` to estimate the **logit** model
- Use `help(crime1)` to get the variable description.

1. Create a figure with two plots alongside each other: Plot 1: the log likelihood function for a Bernoulli distributed random variable together with Plot 2: the numerical derivative (the gradient) of the log likelihood function. Proceed as follows:
   - Simulate a vector x of length $n = 100$ from a Bernoulli distribution using `rbinom(n,1,`θ_0`)` with $\theta_0 = 0.5$.
   - Write a function which takes x and θ as an input and returns the value of the log likelihood function $LLF_n(\theta)$ as an output. Evaluate this function on the grid of $\theta = \{0.01, 0.02, ..., 0.97, 0.98, 0.99\}$. Use the Example 6.1 from the lecture notes a guidance.
   - Create a vector $dx = \theta_i - \theta_{i-1}$ for $i = 2, ..., 99$. Create a vector of the corresponding changes in the log likelihood function $dy = LLF_n(\theta_i) - LLF_n(\theta_{i-1})$. Define the numerical derivative as dy/dx. Use the homework exercise from the computer tutorial 1 as guidance.

   - Create a figure with two subplots next to each other. On the left-hand side panel plot the log likelihood function using the grid of $\theta = \{0.01, 0.02, ..., 0.98, 0.99\}$ as x-axis. Add a vertical line at θ_0. On the left-hand side plot the numerical derivative of the log-likelihood function with the grid of θ as x-axis. Add a vertical line at θ_0 and a horizontal line at zero.

2. Create a figure with two histograms alongside each other: Plot 1: a histogram of the centered maximum likelihood estimator (MLE) of the sample mean of a normally distributed variable and Plot 2: a histogram of the standardized MLE of the sample mean of a normally distributed variable. Proceed as follows:

   - Generate $MC = 5000$ simulation draws of $x_i \sim \mathcal{N}(\mu_0, \sigma_0)$, $i = 1, ..., 100$. Use $\mu_0 = 5$, $\sigma_0 = 10$.
   - For every simulation draw, compute the MLE of the mean $\hat{\mu}_{ML}$. Use Example 6.2 from the lecture notes as guidance.
   - Using the results of Example 6.2 compute the standardization of $\hat{\mu}_{ML}$, which has a limiting standard normal distribution $\mathcal{N}(0, 1)$.
   - Create a figure with two subplots next to each other. On the left-hand side panel plot a histogram of the centered MLE $\hat{\mu}_{ML} - \mu_0$. On the right-hand side panel plot a histogram of the standardized MLE from the previous step. Plot the theoretical density of the standard normal distribution $\mathcal{N}(0, 1)$ on top of each histogram.

3. Estimate a logistic regression with the dependent variable $y = narr86 > 0$ and explanatory variables $x_1 = durat$, $x_2 = pcnv$, $x_3 = ptime86$. Using the library `mfx` calculate the estimated marginal effect for every explanatory variable using the option `atmean = TRUE`. Write down the economic interpretation of the estimated marginal effects, which are significantly different from zero at 5% significance level. Use computer tutorial 8 together with Section 6.2.4 from the lecture notes as guidance. Report your answers in the following table:

	Estimated Marginal Effect	Economic Interpretation
x_1		
x_2		
x_3		