

Mathematical practice final exam 2025

2025-07-02

1.

Find the inverse of the following matrix and verify it using the `all.equal()` function.

$$\begin{pmatrix} 9 & 4 & 12 & 2 \\ 5 & 0 & 7 & 9 \\ 2 & 6 & 8 & 0 \\ 9 & 2 & 9 & 11 \end{pmatrix}$$

2.

Execute the following lines which create two vectors of random integers which are chosen with replacement from the integers $0, 1, \dots, 999$. Both vectors have length 250.

```
xVec <- sample(0:999, 250, replace=T)
yVec <- sample(0:999, 250, replace=T)
```

- (a) Create the vector $(y_2 - x_1, \dots, y_n - x_{n-1})$.
- (b) Pick out the values in `yVec` which are > 600 .
- (c) What are the index positions in `yVec` of the values which are > 600 ?
- (d) Sort the numbers in the vector `xVec` in the order of increasing values in `yVec`.
- (e) Pick out the elements in `yVec` at index positions 1, 4, 7, 10, 13, \dots

3.

For this problem we'll use the (built-in) dataset `state.x77`.

```
data(state)
state.x77 <- as_tibble(state.x77, rownames = 'State')
```

- a. Select all the states having an income less than 4300, and calculate the average income of these states.
- b. Sort the data by income and select the state with the highest income.
- c. Add a variable to the data frame which categorizes the size of population: ≤ 4500 is S, > 4500 is L.
- d. Find out the average income and illiteracy of the two groups of states, distinguishing by whether the states are small or large.

4.

- Write a function to simulate n observations of (X_1, X_2) which follow the uniform distribution over the square $[0, 1] \times [0, 1]$.
- Write a function to calculate the proportion of the observations that the distance between (X_1, X_2) and the nearest edge is less than 0.25, and the proportion of them with the distance to the nearest vertex less than 0.25.

5.

To estimate π with a Monte Carlo simulation, we draw the unit circle inside the unit square, the ratio of the area of the circle to the area of the square will be $\pi/4$. Then shoot K arrows at the square, roughly $K * \pi/4$ should have fallen inside the circle. So if now you shoot N arrows at the square, and M fall inside the circle, you have the following relationship $M = N * \pi/4$. You can thus compute π like so: $\pi = 4 * M/N$. The more arrows N you throw at the square, the better approximation of π you'll have.

```
n <- 10000

set.seed(1)
points <- tibble("x" = runif(n), "y" = runif(n))
```

Now, to know if a point is inside the unit circle, we need to check whether $x^2 + y^2 < 1$. Let's add a new column to the points tibble, called `inside` equal to 1 if the point is inside the unit circle and 0 if not:

```
points <- points |>
  mutate(inside = map2_dbl(.x = x, .y = y, ~ifelse(.x**2 + .y**2 < 1, 1, 0))) |>
  rowid_to_column("N")
```

- Compute the estimation of π at each row, by computing the cumulative sum of the 1's in the `inside` column and dividing that by the current value of `N` column:
- Plot the estimates of π against `N`.

6.

Mortality rates per 100,000 from male suicides for a number of age groups and a number of countries are given in the following data frame.

```
suicrates <- tibble(Country = c('Canada', 'Israel', 'Japan', 'Austria', 'France', 'Germany',
                                'Hungary', 'Italy', 'Netherlands', 'Poland', 'Spain', 'Sweden', 'Switzerland', 'UK', 'USA'),
  Age25.34 = c(22, 9, 22, 29, 16, 28, 48, 7, 8, 26, 4, 28, 22, 10, 20),
  Age35.44 = c(27, 19, 19, 40, 25, 35, 65, 8, 11, 29, 7, 41, 34, 13, 22),
  Age45.54 = c(31, 10, 21, 52, 36, 41, 84, 11, 18, 36, 10, 46, 41, 15, 28),
  Age55.64 = c(34, 14, 31, 53, 47, 49, 81, 18, 20, 32, 16, 51, 50, 17, 33),
  Age65.74 = c(24, 27, 49, 69, 56, 52, 107, 27, 28, 28, 22, 35, 51, 22, 37))
```

- Transform `suicrates` into *long* form.
- Construct side-by-side box plots for the data from different age groups, and comment on what the graphic tells us about the data.

7.

Load the `LaborSupply` dataset from the `{Ecdat}` package and answer the following questions:

```
#data(LaborSupply)
LaborSupply <- read_csv("LaborSupply.csv")

## Rows: 5320 Columns: 7
## -- Column specification -----
## Delimiter: ","
## dbf (7): lnhr, lnwg, kids, age, disab, id, year
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

# create hour and wage variables
labor <- LaborSupply |>
  mutate(hour = exp(lnhr), wage = exp(lnwg), .before = kids) |>
  dplyr::select(-lnhr, -lnwg)
```

- Compute the average annual hours worked and their standard deviations by year.
- What age group worked the most hours in the year 1982?
- Create a variable, `n_years` that equals the number of years an individual stays in the panel. Is the panel balanced?
- Which are the individuals that do not have any kids during the whole period? Create a variable, `no_kids`, that flags these individuals (1 = no kids, 0 = kids)
- Using the `no_kids` variable from before compute the average wage, standard deviation and number of observations in each group for the year 1980 (no kids group vs kids group).