1 Mock Exam 2023

Question 1: Probability (35 marks)

You observe the following sample proportions (iid). F is a random variable which is one for `female` and Y is random variable which refers to `labour-force participation`. Provide formulas for your answers.

Y \ F	1	0	
1	35	45	80
0	5	15	20
	40	60	100

(1)

(a) Define an appropriate sample space Ω. (5 mks.)

(b) Evaluate the prior probabilities π_Y and marginal probabilities π_F. (5 mks.)

(c) Derive the conditional probabilities of F given Y. How do we call these quantities in case we consider F the evidence? (5 mks.)

(d) Derive Bayes rule. (5 mks.)

(e) Evaluate your derivations using the table above for *two different* realisations $\omega \in \Omega$. (5 mks.)

(f) Are Y and F independent, argue why or why not. (5 mks.)

(g) You are getting and updated dataset and go back to your code and find the following function. Explain its purpose and give an example how this function is evaluated and comment. (5 mks.)

```
nyf <- function(y, f) {
    sum(filter(data, Y = y, F = f))
}
```

TURN OVER

Question 2: Statistical Learning (35 marks)

For the following assume the specification

$$Y = f(X) + \varepsilon.$$

(a) Explain, in your own words, the goal of statistical learning. (5 mks.)

(b) What is supervised machine learning and how does it relate to the concept of estimation you know from econometrics? (5 mks.)

(c) Assume Y is a cardinal, quantitative variable. Stating all your steps, show how we can split up the reducible mean squared error $\mathbb{E}[(\widehat{f}-f)^2]$ into variance $\mathbb{E}[(\widehat{f}-\mathbb{E}\widehat{f})^2]$ and squared bias where Bias$(\widehat{f}) = \mathbb{E}\widehat{f} - f$. (5 mks.)

(d) You have many regressors P and do not know which ones to include. Explain, conceptually, how you would approach this problem. Explain the concept of overfitting in this context. (5 mks.)

(e) Let the γ-“norm” for $\gamma = 0, 1, 2$ be (10 mks.)

$$||\beta||_\gamma = \sum_{p=1}^{P} |\beta|^\gamma.$$

For a linear regression model, write down the regularised objective function and explain different variable selection approaches for all cases of gamma (we assume here that $0^0 = 0$).

(f) Your favourite generative AI gives you the following code. (5 mks.)

```
p <- nrow(X)
n <- ncol(X)
b <- Variable(p)
lambda <- 0.5
obj <- Minimize(sum_squares(y - X \%*\% beta) - lambda * sum(x))
sol <- solve(Problem(obj))
coefficients <- sol$getValue(beta)
```

There are 5 mistakes. Find them! **TURN OVER**

Question 3: Classification (30 marks)

We have n i.i.d. realisations of the $(1+P)$-dimensional random vector (Y, X) denoted (y_i, x_i). We consider the logistic regression model:

$$\mathbf{P}(Y = 1|X_i = x) = \Lambda(x^T\beta) \quad (2)$$

where $\Lambda(t) = \frac{1}{1+e^{-t}}$.

(a) Assuming your data is binary, derive the log-odds ratio. What is its shape? (5 mks.)

(b) Derive the log-likelihood of your sample. (10 mks.)

(c) We can think of the logit model as follows: (7 mks.)

$$U_i^* = \delta_1(x) + \xi_i$$

where $Y_i = \mathbf{1}_{\{U_i^*>0\}}$. Show that $\mathbf{P}(Y_i = 1|X = x)$ takes the form in equation (2) and provided that $\xi_i \sim \text{Logistic}_{0,1}$ (standard logistic).

(d) What is the “structural” interpretation of U_i^*? (5 mks.)

(e) Do you think, in general, $\delta_j(x)$ could vary with both choice j and individual i’s observed characteristics, i.e. take the form $x_{ij}\beta_j$? (3 mks.)

END OF EXAM PAPER