Estadística I, Convocatoria extraordinaria, June 23rd Exam
Bachelor's Degrees in ADE, ADE-DER, ADE-INF, ECO, ECO-DER.

EXAM RULES: 1) Use separate booklets for each problem.
2) Perform the calculations with at least two significant decimal places.
3) You cannot leave the exam during the first 30 minutes.
4) You cannot leave the classroom without handing in the exam.

Problem 1 (10 points) The table below shows data about the age and length of service, both measured in years, from 75 employees of a certain company. *Note: blank cells correspond to zero counts.*

	length of service										
age	**1**	**2**	**3**	**4**	**5**	**7**	**9**	**10**	**12**	**15**	**20**
[20,30)	3	6		4							
[30,40)		1	4	2	3						
[40,50)				1	4	13	2	3	1		
[50,60)							2	8	7	4	
[60,70]										2	5

a) (1 point) Indicate, what kind of variables are `age` and `length of service`.

b) (1.5 points) Find the (absolute) marginal distribution of the variable `length of service`. What is the mean length of service for the company's employees?

c) (2 points) Taking into account that the units of the variables are different, compare the variability in `age` with that in `length of service`. Use the following computer output.

```
numSummary(age, statistics=c("mean","sd","quantiles"), quantiles=c(0,.25,.5,.75,1))
         mean        sd      0%   25%  50%  75%   100%   n
age  44.891892   11.730498   20   38   46   53.0  67     75
```

d) (1.5 points) Find the distribution of variable `length of service` for employees aged $[40, 50)$. Roughly, what kind of asymmetry does this distribution have? Answer without plotting any graphs.

e) (2 points) Based on the following computer output, describe the relationship between both variables. Justify your answer.

```
Call:lm(formula=Length of service ~ age, data=Things)

Residuals:
  Min       1Q       Median     3Q      Max
-4.8303   -1.5464   -0.3001   1.4345   6.0651

Coefficients:
             Estimate   Std. Error  t value   Pr(>|t|)
(Intercept)   -8.45427    1.12128     -7.54    1.10e-10 ***
age            0.37315    0.02418     15.44     < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.423 on 72 degrees of freedom
Multiple R-squared: 0.8679, Adjusted R-squared: 0.7647
F-statistic: 238.2 on 1 and 72 DF,  p-value: < 2.2e-16
```

f) (2 points) Based on the computer output from e), write down the linear model that links variable `length of service` with variable `age`. Interpret both coefficients of the model. Based on the model, predict the length of service for an 35-year-old employee. Explain whether this prediction makes sense.

Problem 2 (10 points) The probability density function of a random variable X is

$$f(x) = \begin{cases} x+1, & \text{if } -\frac{1}{2} < x < \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

a) (2.5 points) Find the cumulative distribution function of X.

b) (2.5 points) What is the probability that X is less than 0.

c) (2.5 points) What is the probability that X is greater than $-\frac{3}{2}$ but less than $\frac{1}{4}$.

d) (2.5 points) Calculate the expectation of X.

Problem 3 (10 points) According to a study from the first quarter of 2010, conducted in all branches of a certain bank, the number of fraudulent checks detected by a branch follows a Poisson distribution with a mean of 10 fraudulent checks per day.

a) (3 points) What is the probability that, on a given day, the number of fraudulent checks received by a branch will be less than 3?

b) (4 points) On a given day, we choose a random sample of 200 branches. What is the probability that the mean number of fraudulent checks received by those 200 branches will be greater than 12?

c) (3 points) Now let's assume that we don't know the mean number of fraudulent checks detected by a branch of the bank per day. However, suppose that the daily number of fraudulent checks per branch follows a Poisson distribution with parameter λ, where λ is unknown. If we take a simple random sample of 200 branches, propose an unbiased estimator of λ. Explain your reasoning.