Network Protocol Design and Evaluation

Exercise 8

Stefan Rührup

University of Freiburg
Computer Networks and Telematics
Summer 2009
Task 1

Task 1  Empirical distributions

You observe the remote access to your server for 30 min and count the requests for each minute. The result is the following:

3 3 2 3 1 2 7 0 2 0 1 6 3 1 3 3 4 5 4 7 2 6 4 2 2 3 3 5 4 1

1. Calculate mean, median and quartiles for the number of requests per minute. Calculate the empirical distribution function.

2. Does this data fit to a distribution you know? Show the goodness of fit graphically.

3. Perform a \( \chi^2 \) test of the sample data and a Poisson distribution.
Task 1

- **Data:**
  3 3 2 3 1 2 7 0 2 0 1 6 3 1 3 3 4 5 4 7 2 6 4 2 2 3 3 5 4 1

- **Summary:**
  Min. 0.000
  1st Qu. 1.000
  Median 3.000
  Mean 3.067
  3rd Qu. 4.000
  Max. 7.000
Empirical Distribution Function

Plotted with R using “plot(ecdf(d))”, where d is the data vector.
Empirical and Poisson CDF

Empirical CDF
Poisson CDF, $\lambda=3$
Density Functions

Empirical data histogram
Poisson PDF, $\lambda = 3$
Task 1

- Chi square test:
  - Under a significance level of 0.05 reject the hypothesis, if the result $\chi^2 > 14.07$
  - $\chi^2 = 2.2476$, df = 7, p-value = 0.9449
  - The result means: the test does not indicate a bad fit
> d=c(3,3,2,3,1,2,7,0,2,0,1,6,3,1,3,3,4,5,4,7,2,6,4,2,2,3,3,5,4,1)
> table(d)
 d
0 1 2 3 4 5 6 7
2 4 6 8 4 2 2 2
> tabulate(d)
[1] 4 6 8 4 2 2 2
> f=c(2,tabulate(d))  ← tabulate omits the count for x=0, we add 2 manually
> f
[1] 2 4 6 8 4 2 2 2
> p=f/sum(f)  ← normalize the data
> p
[1] 0.06666667 0.13333333 0.20000000 0.26666667 0.13333333 0.06666667
[7] 0.06666667 0.06666667
> x=c(0,1,2,3,4,5,6)
> q = c(dpois(x, 3),1-sum(dpois(x, 3)))  ← prob for x=0..6 and for x > 6
from the Poisson distribution w. λ=3
> q
[1] 0.04978707 0.14936121 0.22404181 0.22404181 0.16803136 0.10081881
[7] 0.05040941 0.03350854
> chisq.test(f,p=q)  ← how good does the sample data match the probability vector?
Chi-squared test for given probabilities
data:  f
X-squared = 2.2476, df = 7, p-value = 0.9449
>
Task 2

**Task 2** *Arrival Processes*

Assume that the arrival of students in the cafeteria follows a Poisson process and that 3 students arrive on average per minute.

1. You decide to eat there if it is likely (let's say with probability of more than 0.75) that the number of students arriving per minute is less than 5. Will you go there?

2. What is the probability that more than 10 students arrive?

3. Plot the probability mass function and the cumulative distribution function.
Task 2

- Poisson process with $\lambda=3$ arrivals per unit time ($t=1$)
  \[ \Pr_X[X = k] = \frac{\lambda^k}{k!} e^{-\lambda} \]

- Probability that less than 5 students arrive:
  \[ \Pr[X \leq 4] = \sum_{k=0}^{4} \Pr_X[X = k] \approx 0.815 \]

- Probability that more than 10 students arrive:
  \[ \Pr[X > 10] = 1 - \Pr[X \leq 10] \approx 0.00029 \]
Task 2

Probability mass function

Cumulative distribution function

Network Protocol Design and Evaluation
Stefan Rührup, Summer 2009

Computer Networks and Telematics
University of Freiburg