

CSC521 Assignments for Spring 2013

Professor Massimo Di Pierro

May 29, 2013

General Ramarks

- These are individual assignments. You can discuss them with your colleagues but you cannot collaborate on the solution.
- Solutions must include proofs and proofs must include steps
- Formulas must be in electronic format in PDF (not handwritten)
- Your solutions will be cross checked using turnit-in.
- Late assignments are not accepted under any condition.
- All programs must be indented and commented
- The total is 100 points
- You can get extra points by helping reporting errors, bugs and typos.
- Points will be converted to letter grades based on this scale:

A	95-100
A-	92-94
B+	88-91
B	85-87
B-	82-84
C+	78-81
C	75-77
C-	72-74
D+	68-71
D	65-67
D-	62-64
F	0-55

1 Assignment 1 (due April 14, 15 points)

1.1 Problem 1 (7 points)

Consider the following program:

```
import random
def make_data(n=1000):
    table = []
    for i in range(n):
        a,b = random.random(),random.random()
        table.append((a+2.0*b, a*b, a-b))
    return table
```

It returns three columns of numbers. Write a program to compute the mean, variance, and standard deviation of each column. Also compute the correlation and covariance of each two columns.

Analyze the program and determine the theoretical expected values for mean, variance, standard deviations, covariances and correlations.

1.2 Problem 2 (8 points)

Consider a medical insurance company. 20% of the policies are for one person. 80% are family policies. A family can have 2 members, 3 members, 4

members, or 5 members with equal probability. Each person has a 2% probability of filing a claim within a given year. Each claim can range from \$1000 to \$100,000 with equal probability. Assume the company makes no profit and has no operating costs. Also assume that a family policy costs twice as much as an individual policy. How much should the insurance policy cost? Do it numerically and analytically.

2 Assignment 2 (due April 28, 15 points)

2.1 Problem 1 (5 points)

Consider a company that sells mortgage insurance. You insure 10 years, 20 years, and 30 years mortgages. Every month there is a 1% probability that a buyer will default on a mortgage and will not be able to make the remaining payments. How much should you charge per month for each type of mortgage? Assume a 3% annual risk free rate. Provide the code that solves this problem.

2.2 Problem 2 (5 points)

Answer using your own words, each of the following questions (2 points each):

- What are the characteristics of a good random number generator?
- What is the difference between a Poisson distribution and the exponential distribution?
- What are the main differences between the Gaussian distribution and the Pareto distribution?
- When pricing an option, when do you need to use Monte Carlo?
- What are the two ways you know to compute the statistical error on an average?
- What is the meaning and purpose of “resampling”?

3 Assignment 3 (due May 12, 20 points)

3.1 Problem 1

Consider a small shipment company that every day ships N packages where N is random variable following the Poisson distribution with average equal to 100 packages/day. Each package is automatically insured for a random value following the Pareto distribution with $\alpha = 3.0$ and $xm = \$200$. Each package has a 2% probability of being lost. How much capital should the shipping company set aside at the beginning of every year to cover the costs of lost packages, so that the probability of default is less than 3%?

Include in the solution: one page explanation of the solution, the code, an approximate estimate for the solution, a graph depicting the losses as function of time in one random simulated scenario, a plot of the distribution of the results of `simulate_once`.

3.2 Problem 2 (5 points)

Consider the following distributions:

$$p(x) = a/(1 + x^2) \quad x \in [0, \infty] \quad (1)$$

$$p(x) = b\cos(x)/(b\sin(x) + 1) \quad x \in [0, \pi/2] \quad (2)$$

$$p(x) = cx^{M-1}/(x^M - 1)^{1-1/N} \quad x \in [1, 2^{1/M}] \quad (3)$$

For each of them determine the normalization constant a , b and c (M and N are parameters) and use the inversion method to compute a formula to map a uniform random number into a number with the given distribution. Implement each as a Python program.

4 Assignment 4 (due May 26, 15 points)

Solve one of the following two problems:

Submit a single PDF document containing, explanation of the problem, your assumption, your code (indented), documentation for the code, your answers and your explanations. Include diagrams and graphs to explain your results. State all your assumptions for example: “I assume the time interval between two incoming requests follows the ... distribution”, etc.

You will be evaluated on clarity as well as correctness of the results. Assume you are explaining your finding to a manager with an MBA, not to a Math or CS professor.

You must use the bootstrap algorithm to compute the average profit.

File formats other than PDF will not be accepted.

4.1 Problem 1

Consider the following situation. You are a top manager of a large bank and you oversee N traders. Each trader invests a total capital A/N . The daily log-return for each trader is R . R is a random variable that on most days follows a Gaussian distribution with mean $\mu = 0.0005$ and $\sigma = 0.04$ and are not correlated. At random exponential intervals, R follows a Pareto distribution with $\alpha = 1.5$ and $xm = -0.33$. These negative jumps occur in average once every 10 years. The time of the jumps is 100% correlated between traders but the amount of the jumps is not correlated.

What is the total average log return for each trader? What is the average total arithmetic return for all traders?

As a manager you get a 1% annual bonus on the total profit ($A * r$ where r is the total arithmetic return over one year). If the total loss exceeds 20% in one year, you lose your job.

Compute many simulated scenarios for the next one year (`simulate_once`) and determine your average total bonus (zero in those scenarios where you lose your job). Discuss the similarity between this problem and pricing an option using jump diffusion. Discuss the effects of increasing N , decreasing N , increasing A , decreasing A . In our own self interest, what is the best strategy? In your company interest what is the best strategy?

4.2 Problem 2

You run a web service that performs computations on behalf of your users. When a request arrives to your computers, you process it and you deliver an answer. Requests arrive at random intervals with an average of one per minute. Once a request arrives, if you have a computer server available to process it, you assign it and it will complete in a random time given by the Pareto distribution with $\alpha = 2.5$ and $xm = 5$ minutes. If the request arrives and all your servers are busy it will be queued. Each server can deal with one request at the time. One request can be queued but for no longer than 5

minutes. If a request is queued for more than 5 minutes it is dropped. You will have to pay \$10 in penalty for each dropped request.

A server costs \$300 per year. How many servers do you need to maximize the average profit? (assume servers and dropped requests are your only costs)

How many servers do you need so that the 95% VaR is less than \$10,000. What is the average profit in this case?

5 Final Exam

Solve one of the problems listed under assignment 4. Not the one that you have solved already. For CS and Computational Finance students the code must be in C++, Java, or C#, not in Python. For Predictive Analytics students (and other degrees that do not require C++ knowledge) the code can be in Python or SAS.