Assignment A1

MASD 2018

Department of Computer Science
University of Copenhagen

Casper Lisager Frandsen <fsn483@alumni.ku.dk>

Version 1
Due: December 4th, 23:59

Contents

1. Exercise

(a)

To find the pdf of the cdf F, we have to look at the definition, where f is the pdf:

$$F' = f$$

This means we have to find the derivative of:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-(\beta x)^{\alpha}}, & x > 0 \end{cases}$$

With respect to x:

$$f(x) = \begin{cases} \frac{d}{dx}0, & x \leq 0 \\ \frac{d}{dx}1 - e^{-(\beta x)^{\alpha}}, & x > 0 \end{cases}$$

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{\alpha e^{-(\beta x)^{\alpha}}(\beta x)^{\alpha}}{x}, & x > 0 \end{cases}$$

(b)

To find the probability that it stops functioning within the given time interval, we calculate the definite integral between $[0, 4]$, since we know that:

$$F(x \leq 0) = 0$$

We compute the following with Wolfram Alpha:

$$\int_0^4 f(x) = \int_0^4 \frac{\alpha e^{-(\beta x)^{\alpha}}(\beta x)^{\alpha}}{x} dx = F(4) - F(0)$$

$$(1 - e^{-(\frac{4}{4})^2}) - (1 - e^{-(\frac{0}{4})^2}) = 0.632$$

We then subtract this from 1 because that is the reverse case:

$$1 - 0.632 = 0.368$$

For the second part we have to compute the definite integral with boundaries $[5, 10]$:

$$\int_5^{10} f(x) = \int_5^{10} \frac{\alpha e^{-(\beta x)^{\alpha}}(\beta x)^{\alpha}}{x} dx = F(10) - F(5)$$

We compute this with Wolfram Alpha:

$$(1 - e^{-(\frac{10}{4})^2}) - (1 - e^{-(\frac{5}{4})^2}) = 0.208$$

This means that there is a 20.8% chance of the chip breaking in this interval

(c)

To find the median, we have to look at the definition:

$$median := F(m) = \frac{1}{2}$$

We can write this out, and try to solve for m:

$$F(m) = 1 - e^{-(\beta m)^{\alpha}} = \frac{1}{2}$$

$$1 = \frac{1}{2} + e^{-(\beta m)^{\alpha}}$$

$$ln(1) = ln\left(\frac{1}{2}\right) + ln(e^{-(\beta m)^{\alpha}})$$

$$0 - ln\left(\frac{1}{2}\right) = ln(e^{-(\beta m)^{\alpha}})$$

$$ln(2) = -(\beta m)^{\alpha}$$

Since $\alpha, \beta > 0$ we can do the following:

$$\frac{\sqrt[\alpha]{ln(2)}}{\beta} = m$$

(d)

To find $\mathbb{P}(min(X,Y) \leq 4)$ we have to realise that we can calculate the complement simply, as the product of the two definite integrals $\int_0^4 f(x)dx * \int_0^4 f(y)dy$:

$$\int_0^4 f(x)dx = (1 - e^{-(\frac{4}{4})^2}) - (1 - e^{-(\frac{0}{4})^2}) \approx 0.632$$

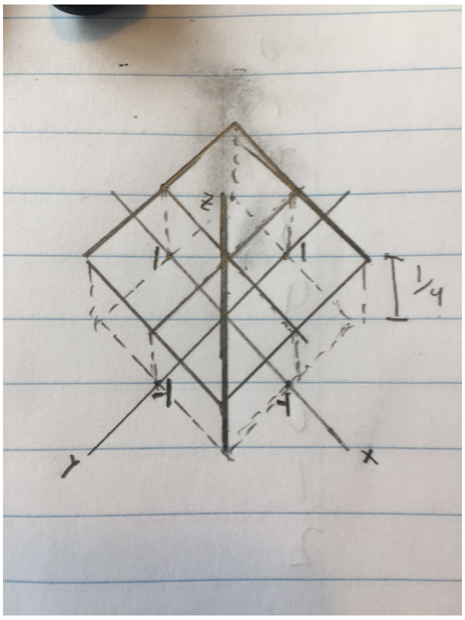
$$\int_0^4 f(y)dy = (1 - e^{-(\frac{4}{6})^3}) - (1 - e^{-(\frac{0}{6})^3}) \approx 0.256$$

$$\mathbb{P}(min(X,Y) \leq 4) = 1 - \left(\int_0^4 f(x)dx * \int_0^4 f(y)dy\right) \approx 1 - 0.632 * 0.256 \approx 0.838$$

This means that there is an 83.8% chance that $\mathbb{P}(min(X,Y) \leq 4)$

2. Exercise

(a)



This may be slightly difficult to read, but the solid lines are either axes or part of the plane resulting from the function. The dotted lines are for ease of viewing, to show that the plane is bounded by the correct values; It should look like a flat plane sitting at $z = \frac{1}{4}$, and be bounded by the values $-1 \leq x, y \leq 1$.

(b)

To compute $\mathbb{P}(X^2 + Y^2 \leq 1)$ we first have to realise that the expression can be visualised as a circle with radius 1. We can then realise that since the two-dimensional integral is equal to the volume under the graph. This means that we simply calculate the volume of a cylinder with a height defined by the pdf given, $\frac{1}{4}$, since the result always stays within the first part of the function:

$$f_{X,Y}(x, y) = \begin{cases} 1/4 & if -1 \leq x, y \leq 1 \\ 0 & otherwise \end{cases}$$

We calculate the volume of the cylinder $\pi * r^2 * h$:

$$\pi * 1^2 * \frac{1}{4} = \frac{\pi}{4}$$

And find that $\mathbb{P}(X^2 + Y^2 \leq 1) = \frac{\pi}{4}$

(c)

This part of the exercise has been completed in the attached Jupyter Notebook named "*A1.ipynb*".
The relevant code can be seen here:

```
def estimator(samples):
    # Input is an series of vectors
    successes = 0
    for i in range (len(samples[0])):
        # Calculating the value for all given vectors,
        # and checking whether they meet the given criteria
        if ((samples[0][i])**2 + (samples[1][i])**2 <= 1):
            successes+=1
    # Normalising the end result to get an estimate of how
    # likely it is for the statement to be true
    return successes/len(samples[0])
```

The estimator is given a range of samples and calculates the average value from it, with the criteria:

$$\mathbb{P}(X^2 + Y^2 \leq 1)$$

The resulting values look as follows:

- estimator at sample size 1000: 0.766
- estimator at sample size 10000: 0.7893
- estimator at sample size 100000: 0.78474
- estimator at sample size 1000000: 0.785553

We can see that this approaches the value $\pi/4 = 0.78539$ as the samples sizes grow larger.

3. Exercise

(a)

To prove that $\hat{\theta}_{stupid}$ is an unbiased estimator, we first have to look at the definition of such:

$$BIAS(\hat{\theta}_{stupid}) = \mathbb{E}\hat{\theta}_{stupid} - \theta_0 = 0$$

We will rewrite this, using the fact that we know the following:

$$\theta_0 = \mathbb{E}X_1$$

$$\mathbb{E}\hat{\theta}_{stupid} = \mathbb{E}\frac{1}{3}\sum_{i=1}^{3}$$

We can now see that we have to prove the following:

$$\mathbb{E}\left(\frac{1}{3}\sum_{i=1}^{3}\right) - \mathbb{E}X_1 = 0$$

Since we know that (X_1, X_2, X_3) are identically distributed, we can say the following:

$$\mathbb{E}X_1 = \mathbb{E}X_n$$

Thus, we can finalise the proof:

$$\frac{\mathbb{E}X_1 + \mathbb{E}X_2 + \mathbb{E}X_3}{3} - \mathbb{E}X_1 = \mathbb{E}X_1 - \mathbb{E}X_1 = 0$$

(b)

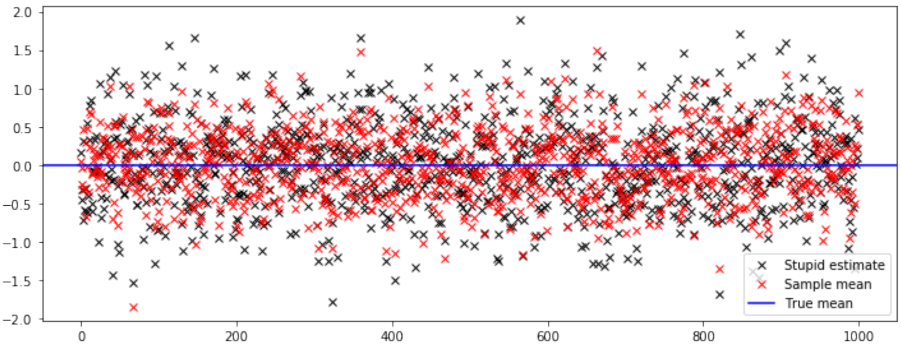
There is quite a bit of code here, but it has been divided into smaller, more readable functions:

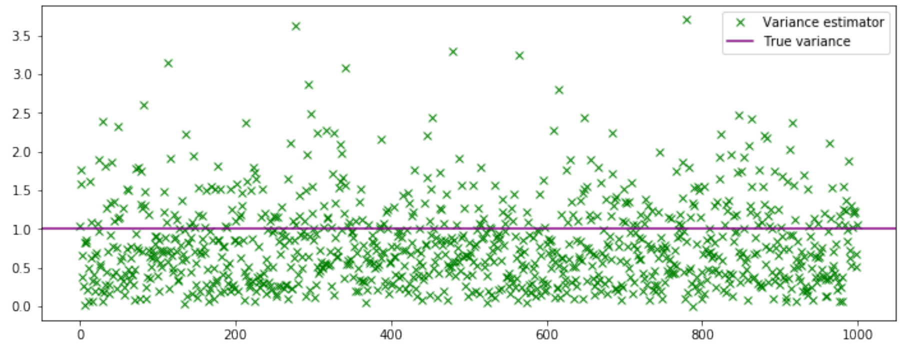
```
# Gaussian sample
def gauss(n):
    return np.random.normal(0,1,n)

# Theta^hat_stupid
def stuEst(samples):
    tempSum = 0
    for i in range(3):
        tempSum += samples[i]
    return 1/3 * tempSum

# Sample Mean
def sampleMean(samples):
    return 1/len(samples) * np.sum(samples)

# Variance Estimator
def varEst(samples):
    return 1/len(samples) * np.sum((samples-sampleMean(samples))**2)

# Calculating and plotting everything n times
def plotAll(distribution, n):
    retVal = np.zeros((3,n))
    for i in range(n):
        temp = gauss(distribution)
        retVal[0,i] = stuEst(temp)
        retVal[1,i] = sampleMean(temp)
        retVal[2,i] = varEst(temp)

    # Lots of irrelevant plotting here, can be seen in the notebook
    return retVal
```

(c)

Looking at the plot seen below there is a noticeably larger variance in the $\hat{\theta}_{stupid}$ estimator, with more extreme values than in the sample mean. This observation is also proven by taking the average of the absolute value of the variances, which gives the following result:

- Stupid estimate average variance: 0.4561300748973738
- Sample mean average variance: 0.36271949341394294





(d)

There is quite a noticeable bias in the variance, as can be seen on the graph above. It tends to be lower than the true variance. This is likely due to the fact that we have used the normal distribution, in which extreme values are extremely uncommon.

(e)

The reason the best estimate is the previous year is because the financial market is an incredibly volatile thing, and the fact that Jeff Bezos' activity may vary quite a bit. Given the current variables, we can say nothing useful about the next year in terms of financial growth because there is an incredibly large amount of variables in play, and every person's growth on the planet is dependant on each others, in one way or another, and that is an impossibly large task to guess just using statistics.

4. Exercise

(a)

To find the *argmax* of the function we first have to find the local extrema, by taking the first derivative, the derivation will be done using Wolfram Alpha. It should be noted that the exercise has not been completed, but the procedure has been explained:

$$\frac{\partial}{\partial\theta}f(x) = \frac{\partial}{\partial\theta}\sum_{i=1}^{n} log(\theta(1-\theta)^{x_i-1}) = \sum_{i=1}^{n}\frac{\theta x_i - 1}{(\theta-1)\theta}$$

We should the solve for $\frac{\partial}{\partial\theta}f(\theta) = 0$ to find critical points. Then, to figure out whether the critical points are the local maximum, we have to find the second-order derivative and solve for the critical points. The derivation will be done using Wolfram Alpha:

$$\frac{\partial^2}{\partial^2\theta}f(\theta) = \frac{\partial^2}{\partial^2\theta}\sum_{i=1}^{n} log(\theta(1-\theta)^{x_i-1}) = \sum_{i=1}^{n}\frac{\theta^2(-x_i)+2\theta-1}{(\theta-1)^2\theta^2}$$

If the result is the < 0 we know that we have found a local maximum. If it is positive we have found a local minimum. If it is $= 0$ we have found a saddle.

$$\sum_{i=1}^{n}\frac{\theta^2(-x_i)+2\theta-1}{(\theta-1)^2\theta^2}, \theta = \text{critical points}$$

(b)