Probability & Statistics: Assignment 7

Due on 20 December, 2018 at 8:30am

Prof. D. Eynard

A. Romanelli / A. Vicini

Question 1

One of the responsibilities of an operating system is to manage and hide the specific hardware implementation of a machine from the user: the operating system should run seamlessly on different type of machines without the user being able to notice. Another responsibility is that of providing an environment for the user application to run in terms of services and resources. The last responsibility is to provide the user with an interface to interact with the system.

Question 2

The distinction between kernel mode and user mode functions as a form of protection for the system because it restricts the commands that a user application (possibly malicious) is able to run to those commands which cannot be harmful.

Question 3

Caches are useful because readings from main memory (or even worse secondary storage systems) are generally much slower than reading from a register in the order of the thousand times. A cache addresses this problem by attempting to "remember" the previously loaded data so that it can be accessed far quicker the next time that the same data needs to be read. Caches are also located physically closer to the CPU, shortening the time needed to access it rather than the main memory.
Because caches are expensive to produce, the amount of data that they are able to store is limited and thus the problem is that the machine must decide what to remember of all the data that is read. Moreover, when multiple processors are present (each with its own cache), the hardware must make sure that these caches are coherent and thus store the same values as each other.
Eliminating the device would be a very bad idea and that would probably be apparent the second the machine is turned off. Since the cache memory is volatile, cutting the power will result in the loss of all the data we had previously stored.

Question 4

a) The first security problem when having multiple jobs running concurrently on the same machine is that data from a user might be accessed and overwritten by another user. Another problem might be that malicious (or incompetent) users may cause the system to block, denying its resources to other users.

b) Under the assumption that the operating system of the time-shared machine is efficient and complex enough to guarantee the fair use of its resources and segregation of user's private data from other users, then it would be possible to achieve the same degree of security, ideally speaking.

Question 5

Problem 2

1. Given the following data: (13.0, 18.5, 16.4, 14.8, 20.8, 19.3, 18.8, 23.1, 25.1, 16.8, 20.4, 17.4, 16.0, 21.7, 15.2, 21.3, 19.4, 17.3, 23.2, 24.9, 15.2, 19.9, 19.1, 18.1, 25.2, 23.1, 15.3, 19.4, 21.5, 16.8, 15.6, 17.6) we computed the sample mean $\overline{x}$ and the standard deviation σ with the following code:

```
samples <- c(13.0, 18.5, ..., 15.6, 17.6)
mu <- mean(samples)
print(mu)
sigma <- sqrt(var(samples))
print(sigma)
```

   Which gives us $\overline{x} = 19.068, \sigma = 3.255$

2. Under the assumption of an underlaying normal distribution and variance $\sigma^2 = 3.12$, we can attempt to find a 95% confidence interval in the following way:

$$[\overline{x} - \delta, \overline{x} + \delta]$$

$$\delta = z^* \frac{\sigma}{\sqrt{n}}$$

   Where the z value for our 95% confidence interval will be $z^* = 1.96$:

$$\delta_{0.95} = 1.96 \cdot \frac{\sqrt{3.12}}{\sqrt{32}} = 1.96 \cdot \frac{1.766}{5.657} = 0.612$$

   Thus we get the resulting :

$$\text{CI}_{0.95} = [\overline{x} - \delta_{0.95}, \overline{x} + \delta_{0.95}] = [19.068 - 0.612, 19.068 + 0.612] = [18.46, 19.68]$$

3. In case of a 99% confidence interval our z^* would be 2.58 instead, which would result in the following margin of error:

$$\delta_{0.99} = 2.58 \cdot \frac{\sqrt{3.12}}{\sqrt{32}} = 0.806$$

   And it's clearly visible that $\delta_{0.95} < \delta_{0.99}$ and thus we'd have a larger confidence interval than before. This comes from the fact that a wider interval grants us more confidence that our sample lies within the interval than not.

   Increasing the sample size n will reduce the margin of error and also tighten the interval.

Problem 3

a) Under the assumption of a biased coin with $\hat{p} = 0.6$, we can compute the probability as the one of a binomial distribution:

$$P(S_n = n \cdot \hat{p}) = \binom{n}{n \cdot \hat{p}} \hat{p}^{n\hat{p}} (1 - \hat{p})^{n(1-\hat{p})}$$

$$P\left(S_n = \frac{3}{5}n\right) = \binom{n}{\frac{3}{5}n} \left(\frac{3}{5}\right)^{\frac{3}{5}n} \left(\frac{2}{5}\right)^{\frac{2}{5}n}$$

For $n = 10$:

$$P(S_{10} = 6) = \binom{10}{6} \left(\frac{3}{5}\right)^{6} \left(\frac{2}{5}\right)^{4} \approx 210 \cdot 0.0466 \cdot 0.0256 = 0.25$$

For $n = 100$:

$$P(S_{100} = 60) = \binom{100}{60} \left(\frac{3}{5}\right)^{60} \left(\frac{2}{5}\right)^{40} \approx 0.081$$

For $n = 1000$:

$$P(S_{1000} = 600) = \binom{1000}{600} \left(\frac{3}{5}\right)^{600} \left(\frac{2}{5}\right)^{400} \approx 0.026$$

Problem 4

Problem 5

To accept or reject H0 we need to compute the z-value:

$$z = \frac{x - \mu}{\sigma/\sqrt{n}} \Rightarrow z = \frac{113.5 - 110}{10/4} = 1.4$$

If the resulting z is greater than what we just computed, then we'll accept H0, else we reject it.

a) A 5% significance level corresponds to $\alpha = 0.05$: $z_{5\%} = \Phi^{-1}(0.95) \overset{?}{>} 1.4$

$$z_{5\%} = 1.65 > 1.4, \text{ thus we accept H0}$$

b) Given $\alpha = 0.10$: $z_{10\%} = \Phi^{-1}(0.90) \overset{?}{>} 1.4$

$$z_{10\%} = 1.28 < 1.4, \text{ thus we reject H0 and accept H1 instead}$$

c) The p-value corresponds to the right-tail-end probability of the null hypothesis H0:

$$p = 1 - \Phi(z) = 1 - \Phi(1.4) \approx 1 - 0.919 = 0.081$$

Problem 6