1 CSE103 Final Practice Problems, set 2

1. **The server**

   Consider a web server that handles a stream of requests. The requests wait for service in a buffer. Assume that this buffer is never empty. The service time varies from request to request.

   Suppose service time for packet i is drawn IID from a fixed but unknown distribtion P whose mean is μ and whose SD is σ, both μ and σ are finite.

   Define T to be the total time for serving the first 1000 packets: $T = \sum_{i=1}^{1000} X_i$.

   - What is the expected value of T? $E[T] = E[\sum_{i=1}^{1000} X_i] = \sum_{i=1}^{1000} E[X_i] = 1000\mu$
   - What is the standard deviation of T? $\sigma[T] = sqrt(1000)\sigma$
   - What is the approximate distribution of T

   Let $A = \frac{1}{1000}T$ be the average service time for the first 1000 packets.

   - What is the expected value of A? $E[A] = \frac{E[T]}{1000} = \mu$
   - What is the standard deviation of A?
     $Var[A] = Var[\frac{T}{1000}] = (\frac{1}{1000})^2 Var[T] = (\frac{1}{1000})^2 1000\sigma^2$
     $\sigma[A] = \frac{\sigma}{\sqrt{1000}}$
   - What is the approximate distribution of A
     $\mathcal{N}(\mu, \frac{\sigma}{\sqrt{(1000)}})$

2. **Statistical tests** **WON'T BE ON THE FINAL**
   The answer to each of the following questions is one of the following tests.

   (a) Z test

   (b) 1 sample t-test

   (c) 2 sample t-test, equal variances

   (d) 2 sample t-test, unequal variances

   (e) Pearson correlation test.

   (f) Spearman correlation test.

   Each of the tests can be either one sided or two sided.

3. **Statistical test** **WON'T BE ON THE FINAL**

   - Suppose the requirement for routers is that they serve 500 packets per second, there is no requirement on the standard deviation.

   - Which test would you use to show that the server's performance does not satisfy the reqirement?

4. **Statistical test****WON'T BE ON THE FINAL**

   - Suppose the requirement for routers is that they serve 500 packets per second, with a standard deviation of 10 packets.

   - Which test would you use to show that the server's performance does not satisfy the reqirement?

5. **Statistical test** **WON'T BE ON THE FINAL**

   - Suppose you are comparing the performance of two servers, the null hypothesis is that their service times have the same mean, but not necessarily the same variance.

   - What test would you use to show that one of the servers is faster than the other.

6. **Statistical test** **WON'T BE ON THE FINAL**

   - Suppose you are comparing the performance of two servers, the null hypothesis is that their service times have the same mean and variance.

   - What test would you use to show that one of the servers is faster than the other.

7. **Statistical test** **WON'T BE ON THE FINAL**

   - Suppose you have two servers and the packets are directed to one of the servers at random. As the packets are different, the service times should be independent, however, you suspect that, if you consider one second at a time, you would see a significant positive correlation.

   - What is the null hypothesis.

   - What test would you use to show that this correlation exists.

8. **Statistical test** **WON'T BE ON THE FINAL**

- Suppose you have 20 servers. You use the appropriate statistical test to check whether the server's performance is better than the requirements. Suppose that you use an threshold $\alpha = 2\%$.

- You run the test and you reject the null hypothesis for server 3.

- What is the level of confidence that you can assign to the statement: "The performance of Server 3 is better than the reqirements" ?

9. **Bounds** Suppose X is a random variable that is never positive. Suppose $E[X] = \mu$. Give a bound on the probability $\mathbf{P}(X < a)$ for $a < E[X]$

   Solution

   Let $Y = -X, E[Y] = E[X]$

   $P(X < a) = P(Y > -a) < \frac{E[Y]}{-a} = \frac{E[X]}{a}$

10. Suppose $X = \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$ and Suppose $E[X] = 0$. Give an upper bound on $\mathbf{P}(X = 4)$ and an upper bound on $\mathbf{P}(X < -2)$.

    Let $Y = X + 4, E[Y] = E[X] + 4 = 4$

    $P(X \geq 4) = P(Y \geq 8) \leq \frac{E[Y]}{8} = \frac{4}{8}$

11. Suppose X, Y are independent random variable with mean μ and standard deviation σ. Give an bound on $\mathbf{P}(|X - Y| > 50\sigma)$

    Let $Z = X - Y, E[Z] = E[X] - E[Y] = 0$

    $Var[X - Y] = Var[X] + Var[Y] = 2\sigma^2$

    $P(|Z - E[Z]| \geq 50\sigma) = \frac{1}{50^2}$