Statistical_Inference1

Laura Bennett

June 3, 2017

Statistical Inference for Data Science: Chapter 5 Exercises

1. If I have a random sample from a population, the sample variance is an estimate of what?

- **The population variance**

2. THe distribution of the sample variance of a random sample from a population is centered at what?

- **The population variance**

3. I keep drawing samples of size n from a population with variance σ^2 and taking their average. I do this thousands of times. If I were to take the variance of the collection of averages, what would it be?

- **In this case, you are seeking the standard error of the mean, a measure of the variability of the mean of a distribution. Thus, the variance of the collection of averages should be** $\frac{\sigma^2}{n}$

4. You get a random sample of n observations from a population and take their average. You would like to estimate the variability of averages of n observations from this population to better understand how precise of an estimate it is. Do you need to repeatedly collect averages to do this?

- **No; you may instead divide the variance by** n **and obtain the standard error of the mean.**

5. A random variable takes the value -4 with probability 0.2 and 1 with probability 0.8. What is the variance of this random variable?

- **The variance of a random variable may be found as** $Var(X) = E(X^2) - E(X)^2$**, where** $E(X) = \sum x_i p(x)$**. In this case,** $E(X) = (-4)(0.2) + (1)(0.8) = 0$ **and** $E(X^2) = (-4)^2(0.2) + (1)^2(0.8) = 4.0$**; therefore,** $Var(X) = 4.0 - 0 = 4.0$

6. If $\bar{X}$ anad $\bar{Y}$ are comprised of n iid random variabls arising from distributions having means μ_x and μ_y respectively and common variance σ^2, what is the variance $\bar{X} - \bar{Y}$?

- **According to the central limit theorem,** $Var(X \pm Y) = Var(X) + Var(Y)$ **IF** X **and** Y **are independent. Given that the sample variance from both** X **and** Y **is** $\frac{\sigma^2}{n}$**,** $Var(\bar{X} - \bar{Y}) = \frac{2\sigma^2}{n}$

7. Let X be a random variable have a standard deviation σ. What can be said about the variance of $\frac{X}{\sigma}$?

- **The variance of a constant times a random variable is given by** $Var(aX) = a^2 Var(X)$**. Thus,** $Var\left(\frac{X}{\sigma}\right) = \frac{\sigma^2}{\sigma^2} = 1$

8. What would be the variance for a random variable with values [2:5] and accompanying probabilities (0.1, 0.2, 0.3, 0.4)?

- ```
  p<-c(.1,.2,.3,.4)
  x<-2:5
  variance <- sum(x^2*p)-(sum(x*p))^2
  print(variance)
  ```

  ```
  ## [1] 1
  ```

9. If you roll ten standard dice, take their average, then repeat this process over and over and construct a histogram, what would be its variance expressed to 3 decimal places?
  ```

- **The variance of a standard die may be calculated as** $Var(X) = \frac{1}{6}\left[1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2\right] - \frac{1}{6^2}\left[1 + 2 + 3 + 4 + 5 + 6\right]^2 = 2.92$**. Since we are essentially looking for the variance of the mean, we may calculate it as** $Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{2.92}{10} = 0.292$