Problem Set 5

Adam Ten Hoeve

Introduction

As an application of some of the properties of expected values, this problem set steps through a proof that the expected value of the random variable that defines sample variance is the population variance, given that the population variance is defined.

For each of these questions, let $X_1, X_2, ...X_n$ be independent, identically distributed random variables with mean μ and variance σ^2.

Please complete the following tasks regarding the data in R. Please generate a solution document in R markdown and upload the .Rmd document and a rendered .doc, .docx, or .pdf document. Please turn in your work on Canvas.

These questions were rendered in R markdown through RStudio (https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf, http://rmarkdown.rstudio.com).

Question 1 (10 points)

Let $X_1, X_2, ...X_n$ be independent, identically distributed random variables with mean μ and variance σ^2, and define the random variable $\bar{X}$ by $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Justify the equality

$$E\left[\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] = E\left[\sum_{i=1}^{n} X_i^2\right] - 2E\left[\sum_{i=1}^{n} \bar{X}X_i\right] + E\left[\sum_{i=1}^{n} \bar{X}^2\right]$$

$$\begin{aligned}
E\left[\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] &= E\Big[\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}X_i + x_i^2\right)\Big] \\
&= E\Big[\sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} 2X_i\bar{X} + \sum_{i=1}^{n}\bar{X}^2\Big] \\
&= E\Big[\sum_{i=1}^{n} X_i^2\Big] - 2E\Big[\sum_{i=1}^{n} X_i\bar{X}\Big] + E\Big[\sum_{i=1}^{n}\bar{X}^2\Big] \checkmark
\end{aligned}$$

Question 2 (10 points)

Let $X_1, X_2, ...X_n$ be independent, identically distributed random variables with mean μ and variance σ^2.

In terms of μ and σ^2, what is the value of $E[X_i^2]$? Note that $Var[X_i] = E[X_i^2] - E[X_i]^2$, while $Var[X_i] = \sigma^2$ and $E[X_i] = \mu$. Please justify your answer.

Confirm numerically that your answer is correct for $X_i \sim exponential(rate = \frac{1}{3})$ which has mean equal to 3 and variance equal to 9.

$$
\begin{aligned}
Var[X_i] &= E[X_i^2] - E[X_i]^2 \\
E[X_i^2] &= Var[X_i] + E[X_i]^2 \\
E[X_i^2] &= \sigma^2 + \mu^2
\end{aligned}
$$

If we have random variables $X_i \sim exp(1/3)$, then we can expect $E[X_i^2] = \sigma^2 + \mu^2 = 9 + (3)^2 = 18$. Let's check if our answer is correct.

```
f2<-function(x){x^2*dexp(x,1/3)}
integrate(f2,0,Inf)
```

```
## 18 with absolute error < 0.00054
```

Question 3 (5 points)

Assuming that $E[X_i^2] = \sigma^2 + \mu^2$ for all i, what is $E\left[\sum_{i=1}^n X_i^2\right]$. Recall that $E\left[\sum_{i=1}^n Y_i\right] = \sum_{i=1}^n E[Y_i]$ for any random variables $Y_i, Y_2 ... Y_n$ with defined means.

$$
\begin{aligned}
E[\sum_{i=1}^n X_i^2] &= \sum_{i=1}^n E[X_i^2] \\
&= \sum_{i=1}^n (\sigma^2 + \mu^2) \\
&= n(\sigma^2 + \mu^2)
\end{aligned}
$$

Question 4 (5 points)

Define the random variable $\bar{X}$ by $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. What is the value of $E\left[\sum_{i=1}^n \bar{X}^2\right]$? Please justify your answer.

Recall that the mean of $\bar{X}$ equals μ and the variance equals $\frac{\sigma^2}{n}$. The fact that $E\left[\sum_{i=1}^n Y_i\right] = \sum_{i=1}^n E[Y_i]$ mentioned above may also be useful. Further, $\bar{X}$ is constant with respect to the index i in the sum.

$$
\begin{aligned}
E\left[\sum_{i=1}^n \bar{X}^2\right] &= \sum_{i=1}^n E[\bar{X}^2] \\
&= \sum_{i=1}^n \left(Var[\bar{X}] + E[X]^2\right) \\
&= \sum_{i=1}^n (\frac{\sigma^2}{n} + \mu^2) \\
&= n(\frac{\sigma^2}{n} + \mu^2)
\end{aligned}
$$

Question 5 (10 points)

Why is

$$E\left[\sum_{i=1}^{n}\bar{X}X_i\right]=E\left[\bar{X}\sum_{i=1}^{n}X_i\right]=E\left[n\bar{X}^2\right]$$

Because $\bar{X}$ is a constant with respect to the index i, we can pull it out of the summation. Then recall $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \rightarrow \sum_{i=1}^{n}X_i=n\bar{X}$. If we can substitute that into our equation, then we get the resulting $E[n\bar{X}^2]$.

Question 6 (5 points)

Assuming that $E\left[\sum_{i=1}^{n}X_i^2\right]=n\left(\sigma^2-\mu^2\right)$, that $E\left[\sum_{i=1}^{n}\bar{X}^2\right]=n\left(\frac{\sigma^2}{n}-\mu^2\right)$, and that $E\left[\sum_{i=1}^{n}\bar{X}X_i\right]=E\left[n\bar{X}^2\right]=n\left(\frac{\sigma^2}{n}-\mu^2\right)$, please simplify $E\left[\sum_{i=1}^{n}X_i^2\right]-2E\left[\sum_{i=1}^{n}\bar{X}X_i\right]+E\left[\sum_{i=1}^{n}\bar{X}^2\right]$.

$$\begin{aligned}
E\left[\sum_{i=1}^{n}X_i^2\right]-2E\left[\sum_{i=1}^{n}\bar{X}X_i\right]+E\left[\sum_{i=1}^{n}\bar{X}^2\right]&=n(\sigma^2+\mu^2)-2n(\frac{\sigma^2}{n}+\mu^2)+n(\frac{\sigma^2}{n}+\mu^2)\\
&=n\sigma^2-2\sigma^2+\sigma^2+n\mu^2-2n\mu^2+n\mu^2\\
&=(n-1)\sigma^2
\end{aligned}$$

Question 7 (5 points)

If $E\left[\sum_{i=1}^{n}X_i^2\right]-2E\left[\sum_{i=1}^{n}\bar{X}X_i\right]+E\left[\sum_{i=1}^{n}\bar{X}^2\right]=(n-1)\sigma^2$, what is the value of $E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]$?

$$\boxed{E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]=\frac{1}{n-1}\left[(n-1)\sigma^2\right]=\sigma^2}$$