Stat474W/574 HW 2

Dr. Islam

Due Jan 27, 2022

Recall the following definitions:

**Definition 1.1**

Given a finite population , the population mean of , denoted by , and the population variance of , denoted by , are defined as follows:

**Definition 1.2**

Given a sample from a population with mean and variance , the sample mean of , denoted by and sample variance of , denoted by are defined as follows:

**Definition 1.3**

* Given two samples and from two populations with unknown but common variance , an estimate of the common variance is given by

where and . The quantity is called the pooled sample variance of the unknown, common variance of the two populations.

# HW 2.1

Regarding the computation of the population variance, under usual notations, the following expressions are popularly used:

The data below refers to the zinc concentration (in mg/ml) in the blood of a population of laboratory rats:

1.76 1.20 1.13 1.71 1.31 1.34 1.47 1.12 1.59 1.03

1. Assign this population data values to a vector , and compute the population mean.
2. Compute the population variance using three expressions provided in (i)-(iii).
3. What is your conclusions based on the computed values in (i)-(iii)?.

#(a)  
X=scan(text="1.76 1.20 1.13 1.71 1.31 1.34 1.47 1.12 1.59 1.03")  
N=length(X)  
mux<-mean(X)  
# (b)  
pop.var1<-sum((X-mux)^2)/N  
pop.var2<-(sum(X^2)-N\*mux^2)/N  
pop.var3<-sum(X^2)/N-mux^2  
c(pop.var1, pop.var2, pop.var3)

[1] 0.059504 0.059504 0.059504

#(c) all three expressions provide identical results.

# Example 2

Regarding the computation of the sample variance, under usual notations, the following expressions are popularly used:

Below is a sample of creative writing score of students in a writing class, after receiving intrinsic motivation:

19.3 21.3 22.6 20.3 17.5 22.2 26.7 29.7

1. Assign this sample data values to a vector , and compute the sample mean.
2. Compute sample variance using three expressions provided in (i)-(iii).
3. What is your conclusions based on the computed values in (i)-(iii)?.

#(a)  
x<-scan(text="19.3 21.3 22.6 20.3 17.5 22.2 26.7 29.7")  
n=length(x)  
xbar<-mean(x)  
# (b)  
samp.var1<-sum((x-xbar)^2)/(n-1)  
samp.var2<-(sum(x^2)-n\*xbar^2)/(n-1)  
samp.var3<-sum(x^2)/(n-1)-(n/(n-1))\*xbar^2  
c(samp.var1, samp.var2, samp.var3)

[1] 15.86857 15.86857 15.86857

#(c) all three expressions provide identical results.

# Example 3

Two samples and , of zinc concentrations (in mg/ml) in the blood for two groups of rats, are given as follows:

x1  
1.15 1.34 1.16 1.53 1.00 1.68 1.55 1.52 1.71 1.59 1.30 1.12 1.42 1.15 1.11  
  
x2  
1.13 1.45 1.23 1.52 1.20 1.34 1.71 1.59 1.55 1.74 1.00 1.15

It is assumed that the two populations the two samples come from have equal, but unknown variance.

1. Compute the value of pooled mean of the two samples given by
2. Compute the value of the pooled sample variance of the two samples (used to estimate the common unknown variance ) given by
3. Compute an estimate of the pooled coefficient of variation given by

x1<-scan(text="1.15 1.34 1.16 1.53 1.00 1.68 1.55 1.52 1.71 1.59 1.30 1.12 1.42 1.15 1.11")  
x2<-scan(text="1.13 1.45 1.23 1.52 1.20 1.34 1.71 1.59 1.55 1.74 1.00 1.15")  
n1<-length(x1)  
n2<-length(x2)  
x1bar=mean(x1)  
x2bar=mean(x2)  
#(a)  
xbarpooled=(n1\*x1bar+n2\*x2bar)/(n1+n2)  
xbarpooled

[1] 1.368148

#(b)  
s1sq<-var(x1)  
s2sq<-var(x2)  
s2pooled<-((n1-1)\*s1sq+(n2-1)\*s2sq)/(n1+n2-2)  
c(s1sq=s1sq,s2sq=s2sq,s2pooled=s2pooled)

s1sq s2sq s2pooled   
0.05382667 0.05928106 0.05622660

#(c)  
coeffv=sqrt(s2pooled)/xbarpooled\*100  
coeffv

[1] 17.33157