HW12

Sumeet Mishra

4/23/2019

1. Trosset Section 15.7 Exercise 4

a) Consider the population of all sister-brother heights. Estimate the proportion of all brothers who are at least 5 feet 10 inches(70 inches).

```
s_ht = c(69,64,65,63,65,62,65,64,66,59,62)
b_ht = c(71,68,66,67,70,71,70,73,72,65,66)
pp=1-pnorm(70,mean(b_ht),sd(b_ht))
sprintf("The proprtion of all brothers who are atleast 70 inches is %f",pp)
```

```
## [1] "The proprtion of all brothers who are atleast 70 inches is 0.356583"
```

Therefore around 35% of the brothers are atleast 70 inches in height.

b) Suppose that Carol is 5 feet 1 inch. Predict her brother's height.

```
carol_ht=61
slope=cor(s_ht,b_ht) * (sd(b_ht)/sd(s_ht))
intercept=mean(b_ht) - slope * mean(s_ht)
predict.yhat=intercept + slope * carol_ht
sprintf("The height of brother when Carol's height is 61 inches is %f",predict.yhat)
```

```
## [1] "The height of brother when Carol's height is 61 inches is 67.227273"
```

Therefore the height of brother when Carol height is 61 inches is 67.227.

c) Consider the population of all sister-brother heights for which the sister is 5 feet 1 inch. Estimate the proportion of these brothers who are at least 5 feet 10 inches.

```
r=cor(s_ht,b_ht)
predict.b=intercept + slope * s_ht
sse=sum((predict.b-b_ht)^2)
pred.error=sqrt(sse/(length(b_ht)-2))
b.pp=1-pnorm(70,predict.yhat,pred.error)
sprintf("The proportion of these brothers who are atleast 70 inches given sister is %f",b.pp)
```

```
## [1] "The proportion of these brothers who are atleast 70 inches given sister is 0.121949"
```

Therefore about 12% of the brothers are atleast 70 inches given height of sister is 61 inches.

2. Trosset Section 15.7 Exercise 8

a) Jill scored 80 points on Test 1. She suggests that her missing score on Test 2 be replaced with her score on Test 1, 80 points. What do you think of this suggestion? What score would you advise the professor to assign?

```
test1.mean=75
test1.sd=10
test2.mean=64
test2.sd=12
n=33
cor=0.5
jill_test1=80
jill_bstar=cor * (test2.sd/test1.sd)
```

```
jill_astar=test2.mean - jill_bstar * test1.mean
jill.predict.y=jill_astar + jill_bstar * jill_test1
print(paste("The predicted test 2 score is",jill.predict.y))
```

```
## [1] "The predicted test 2 score is 67"
```

```
pred_error=test2.sd * sqrt(1-cor^2)
prob=1-pnorm(80,jill.predict.y,pred_error)
sprintf("The probability that Jill scores atleast 80 in test 2 is %f",prob)
```

```
## [1] "The probability that Jill scores atleast 80 in test 2 is 0.105481"
```

Since the probability of Jill scoring atleast 80 in Test 2 is very low, i.e, around 10%, we suggest that she be awarded the predicted score, i.e 67 in test 2.

b) Jack scored 76 points on Test 2, precisely one standard deviation above the Test 2 mean. He suggests that his missing score on Test 1 be replaced with a score of 85 points, precisely one standard deviation above the Test 1 mean. What do you think of this suggestion? What score would you advise the professor to assign?

```
jack_test2=76
jack_bstar=cor * (test1.sd/test2.sd)
jack_astar=test1.mean - jack_bstar * test2.mean
jack.predict.y=jack_astar + jack_bstar * jack_test2
print(paste("The predicted test 1 score is",jack.predict.y))
```

```
## [1] "The predicted test 1 score is 80"
```

```
pred.error=test1.sd * sqrt(1-cor^2)
jack.prob=1-pnorm(85,jack.predict.y,pred.error)
sprintf("The probability Jack scores 85 in Test 1 is %f",jack.prob)
```

```
## [1] "The probability Jack scores 85 in Test 1 is 0.281851"
```

Since the probability of Jack scores 85 in Test 1 is low i.e, around 28%, hence I suggest that he be awarded 80 in test 1.

3.

```
bb.wins=read.csv("baseball-wins.txt",sep="")
#View(bb.wins)
```

a) The probability that a randomly selected team wins at least 84.5 games. You can use either the year 1 wins or year 2 wins to answer this question. Using year1.wins

```
y1.wins=bb.wins$year1.wins
prob.y1.wins=1-pnorm(84.5,mean(y1.wins),sd(y1.wins))
sprintf("The probability that a randomly selected team in year 1 wins atleast 84.5 games is %f",prob.y1
```

```
## [1] "The probability that a randomly selected team in year 1 wins atleast 84.5 games is 0.380871"
```

There is 38% chance that a randomly selected team in year 1 wins atleast 84.5 games.

b) The probability that a team that won 95 games one season wins at least 84.5 games the next season.

```
y2.wins=bb.wins$year2.wins
y1wins=95
corr=cor(y1.wins,y2.wins)
bstar=corr * (sd(y2.wins)/sd(y1.wins))
astar=mean(y2.wins) - bstar * mean(y1.wins)
```

```r
predict.wins=astar + bstar * y1wins
print(paste("The predicted wins is",predict.wins))
```

```
## [1] "The predicted wins is 88.2238032946791"
```

```r
predy1winserror=sd(y2.wins) * sqrt(1-corr^2)
y1winsprob=1-pnorm(84.5,predict.wins,predy1winserror)
sprintf("The probability that a team won 95 games one season wins at least 84.5 games the next season is
```

```
## [1] "The probability that a team won 95 games one season wins at least 84.5 games the next season is
```

The probability that a team won 95 games one season wins at least 84.5 games the next season is around ~ 65%.

c) The probability that a team that won 75 games one season wins at least 84.5 games the next season.

```r
y1_wins=75
corr=cor(y1.wins,y2.wins)
b.star=corr * (sd(y2.wins)/sd(y1.wins))
a.star=mean(y2.wins) - b.star * mean(y1.wins)
predictwins=a.star + b.star * y1_wins
print(paste("The predicted wins is",predictwins))
```

```
## [1] "The predicted wins is 77.8782962937392"
```

```r
prederror.wins=sd(y2.wins) * sqrt(1-corr^2)
y1wins.prob=1-pnorm(84.5,predictwins,prederror.wins)
sprintf("The probability that a team won 75 games one season wins atleast 84.5 games the next season is
```

```
## [1] "The probability that a team won 75 games one season wins atleast 84.5 games the next season is
```

The probability that a team won 75 games in one season wins at least 84.5 games the next season is around ~ 25%.