Homework due Nov 19, 2023 00:37 CST

Make sure that you are using the correct random number generator (RNG) settings by calling the following command:

```
RNGkind("Mersenne-Twister", "Inversion", "Rounding")
```

Smoothing Exercises #1

1/1 point (graded)

Use the data generated in a previous question about men's and women's heights:

```
n = 10000
set.seed(1)
men = rnorm(n,176,7) #height in centimeters
women = rnorm(n,162,7) #height in centimeters
y = c(rep(0,n),rep(1,n))
x = round(c(men,women))
##mix it up
ind = sample(seq(along=y))
y = y[ind]
x = x[ind]
```

Set the seed at 5, `set.seed(5)`, and take a random sample of 250 individuals from the population like this:

```
set.seed(5)
N = 250
ind = sample(length(y),N)
Y = y[ind]
X = x[ind]
```

Use `loess()` to estimate $f(x) = E(Y|X = x)$ using the default parameters. What is the predicted $f(168)$?

0.5480233

✔ Answer: 0.5480233

0.5480233

Explanation

```
fit=loess(Y~X)
predict(fit,newdata=data.frame(X=168))

##Here is a plot
xs = seq(160,178)
Pr =sapply(xs,function(x0) mean(Y[X==x0]))
plot(xs,Pr)
fitted=predict(fit,newdata=data.frame(X=xs))
lines(xs,fitted)
```

Submit

You have used 1 of 5 attempts

Answers are displayed within the problem

Smoothing Exercises #2

1/1 point (graded)

The loess estimate above is a random variable thus we can estimate its standard deviation. Use Monte Carlo simulation to compute the standard deviation of your estimate of $f(168)$ (remember, we have the *entire population*).

Set the seed to 5 with `set.seed(5)` and perform 1000 simulations of the computations performed in question #1. Report the the population standard deviation of the loess based estimate.

0.05755689

✔ Answer: .05487204

0.05755689

Explanation

```
set.seed(5)
nsims = 1000
estimate = rep(NA, nsims)
for (i in 1:nsims) {
  N = 250
  ind = sample(length(y),N)
  Y = y[ind]
  X = x[ind]
  fit = loess(Y~X)
  estimate[i] = predict(fit,newdata=data.frame(X=168))
}
popsd(estimate)
```

Submit

Answers are displayed within the problem