Stat 605 – Spatial Statistics – Spring 2022

Homework 3. Due: Friday, January 28, midnight.

1. Use the following R code to simulate and plot a data set consisting of $n = 60$ values from the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$, where the ϵ's are independent $N(0, \sigma^2)$ random variables.

```
library(geoR)
set.seed(32123) # for reproducibility
n <- 60
lons <- runif( n, min=90, max=110 )
lats <- runif( n, min=35, max=45 )
y <- 30.0 + 1.3*(lons-100) + 1.5*(lats-40) + rnorm(n,mean=0,sd=1.6)
mydata <- as.geodata( cbind(lons,lats,y) )
plot(mydata)
```

   (a) Include the resulting plot in your write-up.

   (b) What are the values of β_0, β_1, β_2, and σ^2 that are being used to simulate these data?

   (c) Fit a linear model with independent errors using the R function `lm`, for example by typing

```
myfit <- lm( y ~ lons + lats )
summary(myfit)
```

   (d) State the estimated regression function and the estimate of σ^2. (Be sure to include the R output.)

   (e) Construct 95% confidence intervals for the β's.

   (f) Make a table with the values of the β's, the estimates of the β's, and 95% CIs for the β's. Comment briefly; for example, are the $\hat{\beta}$'s "close to" the true β's? Do the β's lie in their respective 95% CI's?

   (g) Perhaps $Y(\boldsymbol{s})$ is the weight of observed dog fur, in nanograms, at $\boldsymbol{s}$. Interpret β_0 in terms of $\mathbb{E}(Y)$. Explain why it doesn't really make sense to interpret β_0. (I'm looking for two reasons. Three, if you include "it's silly to talk about dog fur.")

   (h) Interpret β_1 and β_2 in terms of $\mathbb{E}(Y)$. (β_1 is the change in the mean response if ...if what? — If you have questions about I'm looking for, please ask.)

2. Let $\Sigma = \begin{bmatrix} 5 & -3 \\ -3 & 1 \end{bmatrix}$.

   (a) Show that Σ cannot be a valid variance-covariance matrix for $(Y_1, Y_2)^T$, by assuming it is and finding the correlation of Y_1 and Y_2.

   (b) Show that Σ cannot be a valid variance-covariance matrix by assuming it is and finding $\text{Var}(Y_1 + 2Y_2)$.

   In each case, be sure to say why your calculation has shown that Σ can't be a variance-covariance matrix.

Continued on next page.

3. Consider the covariogram $C(h) = \begin{cases} 8 & \text{if } h = 0 \\ 5\exp(-h^2/16) & \text{if } h > 0 \end{cases}$

   (a) Sketch $C(h)$, for example, using this R code:

```
hh <- seq( 0,10, length=200 )
curve( 5*exp(-x^2/16), xlab="lag h", ylab="C(h)",
      from=0.0, to=10)
abline( v = 0 ) # include y-axis on the plot
abline( h = 0 ) # include x-axis on the plot
```

   (b) Calculate the corresponding semivariogram. (Be sure to show your work.) You need to be careful about your calculations at $h = 0$.

   (c) What kind of (semi)variogram is this, and what are its parameters?

   (d) What is the nugget? the partial sill? the sill? the range (or effective range)? $\mathrm{Var}(Y(\boldsymbol{s}))$?

4. Consider the semivariogram $\gamma(h) = \begin{cases} 0 & \text{if } h = 0 \\ 3 + 7(1.5(h/6) - 0.5(h/6)^3) & \text{if } 0 < h < 6 \\ 10 & \text{if } h \geq 6 \end{cases}$

   (a) Use R to sketch $\gamma(h)$ for a suitable range of h.

   (b) Find the corresponding covariogram, $C(h)$.

   (c) What kind of (semi)variogram is this, and what are its parameters?

   (d) What is the nugget? the partial sill? the sill? the range? $\mathrm{Var}(Y(\boldsymbol{s}))$?

5. Let X be a discrete random variable such that

x	-2	-1	1	2
$P(X = x)$	.4	.1	.1	.4

   Define $Y = X^2$.

   (a) Show that $\mathrm{cov}(X, Y) = 0$. (To do this, you will need to calculate $\mathbb{E}(X)$, $\mathbb{E}(Y)$, and $\mathbb{E}(XY)$, and use the formula for covariance from page 25 of the lecture notes.)
   - This part of the question is asking you to show that X and Y are uncorrelated.

   (b) Show that X and Y are dependent. (E.g., calculate $P(X = 1, Y = 4)$ and $P(X = 1)P(Y = 4)$; if X and Y are independent, these two values – among others – must be equal.)

   This problem shows a pair of random variables, X and Y, that are uncorrelated yet dependent. (In this case, if I tell you the value of X, you know everything about Y; and if I tell you the value of Y, e.g. $Y = 1$, you know a great deal about the value X. This shows that X and Y aren't independent.)