Name: ____________________________

Directions: **Work only on this sheet** (on both sides, if needed); do not turn in any supplementary sheets of paper. There is actually plenty of room for your answers, as long as you organize yourself BEFORE starting writing. SHOW YOUR WORK. Some problems may require answers to be in the form of numerical expression. An example is

$$2/7 \cdot 1.39 + \sqrt{29.002} + \int_0^1 0.82t \, dt + (1,2) \left( \begin{array}{cc} 1 & 4 \\ 6 & 1 \end{array} \right) \left( \begin{array}{c} 6 \\ 6 \end{array} \right)$$

No variables are allowed in numerical expressions, nor are "sigma sum" series.

**1.** (15) Consider the quadratic model for mean wait time versus backoff probability in the ALOHA model, p.155-156 and Equation (6.8). Suppose we have just four b values, instead of 100. Give a numerical expression for Q.

**2.** (35) The code below reads in a file, **data.txt**, with the header record

```
"age", "weight", "systolic blood pressure", "height"
```

and then does the regression analysis.

Suppose we wish to estimate $\beta$ in the model

$$\text{mean weight} = \beta_0 + \beta_1 \text{height} + \beta_2 \text{age}$$

Fill in the blanks in the code:

```
dt <- ____________(____________________________________)
regr <- lm(__________________________________________________)
cvmat <- _______________(regr)
cat("the estimated value of beta2-beta0 is",
   ____________________________________________________)
cat("the estimated variance of beta2 - beta0 is",
   _______________________ cvmat _________________________)
# calculate the matrix Q
q <- cbind(______________________________________________)
```

**3.** (30) Say we have a random sample $X_1, ..., X_n$ from a uniform distribution on 1,2,...,c, with c unknown and to be estimated. As shown in Section 4.4.3, the MLE is $\max_i X_i$. Fill in the blanks in the following simulation code:

```
sim <- function(c,n,nreps) {
   mle <- vector(length(nreps))
   for (i in 1:nreps):
      x <- sample(________________________________)
      mle[i] <- _____________(________________________________)
  cat("bias = ",___________________ - ________________________)
}
```

**4.** (20) Say we have a random sample $X_1, ..., X_n$ from a population with mean $\mu$ and variance $\sigma^2$. The usual estimator of $\mu$ is the sample mean $\bar{X}$, but here we will use what is called a *shrinkage estimator*: Our estimate of $\mu$ will be $0.9\bar{X}$. Find the mean squared error of this estimator, and give an inequality (you don't have to algebraically simplify it) that shows under what circumstances $0.9\bar{X}$ is better than $\bar{X}$. (Strong advice: Do NOT "reinvent the wheel." Make use of what we have already derived.)

**Solutions:**

**1** Let d = 0.7/4. Then

$$Q = \left( \begin{array}{ccc} 1 & 0.2 & (0.2+d)^2 \\ 1 & 0.2 & (0.2+2d)^2 \\ 1 & 0.2 & (0.2+3d)^2 \\ 1 & 0.2 & (0.2+4d)^2 \end{array} \right)$$

**2.**

```
dt <- read.table("data.txt",header=T)
regr <- lm(dt[,2] ~ dt[,4]+dt[,1])
cvmat <- vcov(regr)
cat("the estimated value of beta2-beta0 is",
   dt$coef[3] -dt$coef[1],"\n")
cat("the estimated variance of beta2hat - beta0hat is",
   t(c(-1,0,1)) %*% cvmat %*% c(-1,0,1),"\n")
# calculate the matrix Q
q <- cbind(1,nrow(dt),dt[,4],dt[,1])  # note recycling
```

**3.**

```
sim <- function(c,n,nreps) {
   mle <- vector(length(nreps))
   for (i in 1:nreps):
      x <- sample(1:c,n,replace=T)
      mle[i] <- max(x)
  cat("bias = ",mean(mle) - c)
}
```

**4.**

$$Var(0.9\bar{X}) = 0.81\sigma^2/n$$

$$\text{bias}(0.9\bar{X}) = 0.9\mu - \mu$$

$$Var(\bar{X}) = \sigma^2/n$$

$$\text{bias}(\bar{X}) = \mu - \mu = 0$$

So, $0.9\bar{X}$ is better than $\bar{X}$ if

$$0.81 \cdot \frac{\sigma^2}{n} + 0.01\mu^2 < \frac{\sigma^2}{n}$$

or (optionally simplifying)

$$\mu < \sqrt{\frac{19}{n}} \, \sigma \tag{1}$$