MthStat 465, Spring 2005, Homework Number 6

1. Suppose that R_1, R_2, R_3 are independent random variables with the same expected value, μ, and the same standard deviation σ. Show that $a_1R_1 + a_2R_2 + a_3R_3$ is an unbiased estimate for μ if and only if $a_1 + a_2 + a_3 = 1$.

   Note that

$$\begin{aligned}\mathrm{E[a_1R_1 + a_2R_2 + a_3R_3]} &= a_1\mathrm{E[R_1] + a_2E[R_2] + a_3E[R_3]}\\ &= a_1\mu + a_2\mu + a_3\mu\\ &= (a_1 + a_2 + a_3)\mu\end{aligned}$$

and this last expression will always equal μ if and only if $a_1 + a_2 + a_3 = 1$.

Compute the variance of this estimate if $a_1 = 1/2$, $a_2 = 1/3$, and $a_3 = 1/6$.

$$\begin{aligned}\mathrm{Var[a_1R_1 + a_2R_2 + a_3R_3]} &= \mathrm{Var[a_1R_1] + Var[a_2R_2] + Var[a_3R_3]}\\ &\quad \text{since the } R_j \text{ are independent}\\ &= a_1^2\mathrm{Var[R_1] + a_2^2Var[R_2] + a_3^2Var[R_3]}\\ &= (1/4)\sigma^2 + (1/9)\sigma^2 + (1/36)\sigma^2\\ &= (7/18)\sigma^2.\end{aligned}$$

Show that $b_1R_1^2 + b_2R_2^2 + b_3R_3^2$ is an unbiased estimate of $\sigma^2 + \mu^2$ if $b_1 + b_2 + b_3 = 1$.

Recall that for any random variable $\mathrm{Var[R] = E[R^2] + (E[R])^2}$. Therefore

$$\begin{aligned}\mathrm{E[b_1R_1^2 + b_2R_2^2 + b_3R_3^2]} &= b_1\mathrm{E[R_1^2] + b_2E[R_2^2] + b_3E[R_3^2]}\\ &= b_1(\sigma^2 + \mu^2) + b_2(\sigma^2 + \mu^2) + b_3(\sigma^2 + \mu^2)\\ &= (b_1 + b_2 + b_3)(\sigma^2 + \mu^2)\\ &= (\sigma^2 + \mu^2).\end{aligned}$$

Show that $c_1R_1^2 + c_2R_2^2 + c_3R_3^2 + c_4R_1R_2 + c_5R_1R_3 + c_6R_2R_3$ is an unbiased estimate of σ^2 if and only if $c_1 + c_2 + c_3 = 1$ and $c_4 + c_5 + c_6 = -1$.

Recall that if R_j and R_k are independent then $\mathrm{E[R_jR_k] = E[R_j]E[R_k]}$. Therefore

$$\begin{aligned}&\mathrm{E[c_1R_1^2 + c_2R_2^2 + c_3R_3^2 + c_4R_1R_2 + c_5R_1R_3 + c_6R_2R_3]}\\ &c_1\mathrm{E[R_1^2] + c_2E[R_2^2] + c_3E[R_3^2] + c_4E[R_1R_2] + c_5E[R_1R_3] + c_6E[R_2R_3]}\\ &c_1(\sigma^2 + \mu^2) + c_2(\sigma^2 + \mu^2) + c_3(\sigma^2 + \mu^2) + c_4\mu^2 + c_5\mu^2 + c_6\mu^2\\ &(c_1 + c_2 + c_3)(\sigma^2 + \mu^2) + (c_4 + c_5 + c_6)\mu^2\\ =\ &(c_1 + c_2 + c_3)\sigma^2 + (c_1 + c_2 + c_3 + c_4 + c_5 + c_6)\mu^2,\end{aligned}$$

and the only way that this last expression can equal σ^2 for every σ and μ is if

$$\begin{aligned}c_1 + c_2 + c_3 &= 1\\ c_1 + c_2 + c_3 + c_4 + c_5 + c_6f &= 0.\end{aligned}$$

By subtracting the first equation of this system from the second we get the equivalent set of equations

$$\begin{aligned}c_1 + c_2 + c_3 &= 1\\ c_4 + c_5 + c_6f &= -1,\end{aligned}$$

as required.

2. Devise at test to distinguish the following two possibilities:
   - A coin has probability 0.25 of coming up heads on any one toss, and all tosses of the coin are independent;
   - A coin has probability 0.65 of coming up heads on any one toss, and all tosses of the coin are independent.

   Your test should be right at least 90% of the time in either case. Check your results against the formulae derived in Lecture 19. Do not simply "plug in" to the formulae we derived there. Verify that your tests meet the objectives, and adjust them if need be.

   Let H_N be the number of heads in N tosses of the coin. Under either hypothesis H_N has a binomial distribution. We will decide for the 0.25 coin if $H_N \leq C$ and for the 0.65 coin if $H_N > C$. We have to find N and C so that $\Pr_{0.25}(H_N \leq C) \geq 0.90$ and $\Pr_{0.65}(H_N > C) \geq 0.90$.

   If we apply the DeMoivre-Laplace Theorem we get

$$\Pr_{0.25}(H_N \leq C) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(C-0.25N)/sqrt(0.25)(0.75)N} e^{-u^2/2}\, du.$$

   Since we know that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1.2816} \approx 0.9000084999$$

   a good approximation is to require that

$$\frac{C - 0.25N}{\sqrt{(0.25)(0.75)N}} = 1.2816. \tag{1}$$

   We also know that

$$\Pr_{0.25}(H_N > C) \approx \frac{1}{\sqrt{2\pi}} \int_{(C-0.65N)/sqrt(0.65)(0.35)N}^{\infty} e^{-u^2/2}\, du,$$

   and

$$\frac{1}{\sqrt{2\pi}} \int_{-1.2816}^{\infty} \approx 0.9000084999$$

   so another good approximation is that

$$\frac{C - 0.65N}{\sqrt{(0.65)(0.35)N}} = -1.2816. \tag{2}$$

   Solving (1) and (2) together we get $N \approx 8.5$ and $C \approx 3.7$, so we take $N = 9$ and $C = 3$ and see if these give us the right probabilities:

$$\begin{aligned}
\Pr_{0.25}(H_9 \leq 3) &= \sum_{k=0}^{3} \binom{9}{k} (0.25)^k (0.75)^{9-k} \\
&< 0.84, \\
\Pr_{0.65}(H_9 > 3) &= \sum_{k=4}^{9} \binom{9}{k} (0.65)^k (0.35)^{9-k} \\
&> 0.94.
\end{aligned}$$

Increasing C to 4 does not work (what happens?) but increasing C to 4 and N to 10 does work:

$$\begin{aligned}
\Pr_{0.25}(H_{10} \le 4) &= \sum_{k=0}^{4} \begin{pmatrix} 10 \\ k \end{pmatrix} (0.25)^k (0.75)^{10-k} \\
&> 0.921, \\
\Pr_{0.65}(H_9 > 3) &= \sum_{k=5}^{10} \begin{pmatrix} 10 \\ k \end{pmatrix} (0.65)^k (0.35)^{10-k} \\
&> 0.905.
\end{aligned}$$

3. Reformulate the first problem on the first homework assignment in the context of linear models. Set up and solve parts 7, 9 and 10 using matrix methods. You may use a calculator to compute the matrix inverses, but your solution should demontrate that your matrix inverses are correct. Your answers should be the same as you got earlier assuming those answers were correct. Assuming that the errors followed a standard normal distribution, give an unbiased estimate of the variance of the errors for each of your three models.

I will illustrate the method with five pairs of heights and weights instead of fifty. Suppose the height-weight pairs are $(68, 159)$, $(62, 110)$, $(64, 97)$ $(65, 102)$ and $(59, 130)$, where height is in inches and weight is in pounds.

For the linear fit, $Weight = A_1 + A_2 * Height$.

$$\vec{Y} = \begin{bmatrix} 159 \\ 110 \\ 97 \\ 102 \\ 130 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 68 \\ 1 & 62 \\ 1 & 64 \\ 1 & 65 \\ 1 & 59 \end{bmatrix}.$$

The estimate of the vector $[A_1, A_2]^t$ is

$$\begin{aligned}
\vec{A} &= (B^t B)^{-1} B^t \vec{Y} \\
&= \begin{bmatrix} 5 & 318 \\ 318 & 20270 \end{bmatrix}^{-1} \begin{bmatrix} 598 \\ 38140 \end{bmatrix} \\
&\approx \begin{bmatrix} -31.24 \\ 2.37 \end{bmatrix}
\end{aligned}$$

so $A_1 \approx -31.24$ and $A_2 \approx 2.37$.

The vector of differences between the observed weights and the predicted weights is

$$\vec{Y} - B\vec{A} = \begin{bmatrix} 28.96 \\ -5.81 \\ -23.55 \\ -20.92 \\ 21.31 \end{bmatrix}$$

and the square of the length of this vector is about 71775, so an unbiased estimate of σ^2 is about $71775/(5-2) \approx 23925$. Note that the corresponding value of σ is about 154.67.

Now we consider a quadratic model. This changes B by adding a column of the square of the weights, and adds a component to $\vec{A}$ corresponding to the coefficient of the quadratic term in our curve of best fit.

$$\vec{Y} = \begin{bmatrix} 159 \\ 110 \\ 97 \\ 102 \\ 130 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 68 & 68^2 \\ 1 & 62 & 62^2 \\ 1 & 64 & 64^2 \\ 1 & 65 & 65^2 \\ 1 & 59 & 59^2 \end{bmatrix}.$$

Now

$$\begin{aligned} \vec{A} &= (B^tB)^{-1}B^t\vec{Y} \\ &= \begin{bmatrix} 5 & 318 & 20270 \\ 318 & 20270 & 1294908 \\ 20270 & 1294908 & 82902914 \end{bmatrix}^{-1} \begin{bmatrix} 598 \\ 38140 \\ 2438848 \end{bmatrix} \\ &\approx \begin{bmatrix} -9066.27 \\ -285.03 \\ 2.26 \end{bmatrix}. \end{aligned}$$

Notice how small the quadratic coefficient is with respect to the other coefficients.

The vector of differences between the observed weights and the predicted weights is

$$\vec{Y} - B\vec{A} = \begin{bmatrix} 2.63 \\ 9.95 \\ -3.71 \\ -5.83 \\ -3.04 \end{bmatrix}$$

and the square of the length of this vector is about 162.84, so an unbiased estimate of σ^2 is about $162.84/(5-3) \approx 81.42$. Note that the corresponding value of σ is about 9.02.

Finally we consider fitting a cubic function. This changes B by adding a column of the cubes of the weights to the previous B, and adds a component to $\vec{A}$ corresponding to the coefficient of the cubic term in our curve of best fit.

$$\vec{Y} = \begin{bmatrix} 159 \\ 110 \\ 97 \\ 102 \\ 130 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 68 & 68^2 & 68^3 \\ 1 & 62 & 62^2 & 62^3 \\ 1 & 64 & 64^2 & 64^3 \\ 1 & 65 & 65^2 & 65^3 \\ 1 & 59 & 59^2 & 59^3 \end{bmatrix}.$$

Now

$$\begin{aligned}\vec{A} &= (B^tB)^{-1}B^t\vec{Y}\\ &= \begin{bmatrix} 5 & 318 & 20270 & 1294908\\ 318 & 20270 & 1294908 & 82902914\\ 20270 & 1294908 & 82902914 & 5319023148\\ 1294908 & 82902914 & 5319023148 & 341986619210\end{bmatrix}^{-1}\begin{bmatrix}598\\38140\\2438848\\156349756\end{bmatrix}\\ &\approx \begin{bmatrix}-77777.31974\\3826.274595\\-62.51416161\\.3397085610\end{bmatrix}.\end{aligned}$$

Notice how small the cubic coefficient is with respect to the other coefficients, but remember how big the cubes of the heights are.

The vector of differences between the observed weights and the predicted weights is

$$\vec{Y} - B\vec{A} = \begin{bmatrix}-0.11\\0.67\\-1.81\\1.34\\-.09\end{bmatrix}$$

and the square of the length of this vector is about 5.54, so an unbiased estimate of σ^2 is about $5.544/(5-4) = 5.54$. Note that the corresponding value of σ is about 2.35.

Don't be fooled by how small this variance is. You could fit a degree four polynomial exactly through the data and conclude that the variance is 0.