STA 410/2102F, Fall 2007: In-Class Test
SOLUTIONS

1. [5 points] Write (with full explanation) the number -19.25 in double-precision floating-point format.

Solution: $-19.25 = -(16 + 2 + 1 + \frac{1}{4}) = -(2^4 + 2^1 + 2^0 + 2^{-2}) = -2^4(2^0 + 2^{-3} + 2^{-4} + 2^{-6}) = (-1)^1\, 2^4\, 1.001101 = (-1)^1\, 2^{1027-1023}\, 1.001101.$

2. [8 points] Determine (with full explanation) what value R will return if given the following expression: 1 + 2∧(–10) + 2∧50 – 2∧50.

Solution: $1 + 2^{-10} = (-1)^0\, 2^0\, 1.0000000001$ *(base 2).*

Then, adding 2^{50} *gives* $2^0\, 1.0000000001 + 2^{50}\, 1.0 = 2^{50}(2^{-50}\, 1.0000000001 + 1.0) = 2^{50}(0.[49\text{ zeroes}]1[9\text{ zeroes}]1 + 1.0) = 2^{50}\, 1.[49\text{ zeroes}]1[9\text{ zeroes}]1$, *which truncates to* $2^{50}\, 1.[49\text{ zeroes}]1$ *since R only holds 52 decimal points.*

Then, subtracting 2^{50} *gives* $2^{50}\, 1.[49\text{ zeroes}]1 - 2^{50}\, 1.0 = 2^{50}(1.[49\text{ zeroes}]1 - 1.0) = 2^{50}(0.[49\text{ zeroes}]1) = 2^{50}(2^{-50}\, 1.0) = 2^0\, 1.0$, *which equals 1.*

So, the answer is 1.

3. [7 points] Write a complete R program to estimate the quantity $\mathbf{E}[Y/(1 + Z^4)]$, where $Y \sim \text{Poisson}(5)$ and $Z \sim \text{Normal}(0, 1)$, by a Monte Carlo simulation, using 1000 replications. Your program should output both the estimate and its standard error. [Hint: Recall the R commands "rpois(n, lambda)" to generate n independent Poisson(lambda) random variables, and "rnorm(n)" to generate n independent standard normal random variables.]

Solution:

```
n = 1000
yvec = rpois(n,5)
zvec = rnorm(n)
estimate = mean(yvec/(1+zvec∧4))
se = sd(yvec/(1+zvec∧4)) / sqrt(n)
cat("estimate equals", estimate, "\n")
cat("standard error equals", se, "\n")
```

4. Consider the function $g(x) = x^2 - 5$. Suppose we begin with the initial interval $[0, 4]$. [You may assume that $g(0) < 0$ and $g(4) > 0$. And, in what follows, you do <u>not</u> need to write a computer program, you just need to explain what intervals will arise.]

(a) [5 points] Present (with full explanation) the next <u>two</u> intervals computed by the Bisection Method.

Solution: *First midpoint = (0+4)/2 = 2.*

And, $g(2) = 2^2 - 5 = 4 - 5 = -1 < 0$, *but* $g(4) > 0$.

So, the first new interval is $[2, 4]$.

Then, second midpoint = (2+4)/2 = 3.

And, $g(3) = 3^2 - 5 = 9 - 5 = 4 > 0$, *but* $g(2) > 0$.

So, the second new interval is $[2, 3]$.

(b) [5 points] Present (with full explanation) the next <u>one</u> interval computed by the False Position (Safe Bisection) Method.

Solution: *The line joining* $(0, g(0))$ *to* $(4, g(4))$, *i.e. joining* $(0, -5)$ *to* $(4, 11)$, *is the line* $y = 4x - 5$.

Its root is the value of x *such that* $4x - 5 = 0$, *i.e.* $x = 5/4$.

$g(5/4) = (5/4)^2 - 5 = (25/16) - 5 < 0$, *but* $g(4) > 0$.

So, the new interval is $[5/4,\ 4]$.

5. [10 points] Suppose it is believed that $Y = (X + \beta)^2 + \text{error}$, with β unknown. Assume that `x` and `y` are two vectors in R, each of length `n`, corresponding to the observed `x` and `y` values. Write a complete R program which runs the Newton-Raphson Method for 100 iterations, with initial value 2, to compute the least-squares estimate of β. [Hint: before writing the program, you may need to compute some derivatives, etc.; be sure to explain all such computations.]

Solution: *The squared-error function is* $f(\beta) = \sum_{i=1}^{n} (y_i - (x_i + \beta)^2)^2$.

Its derivative is $g(\beta) \equiv f'(\beta) = -4 \sum_{i=1}^{n} (y_i - (x_i + \beta)^2)\,(x_i + \beta)$. *Then, the derivative of* g *is* $g'(\beta) = -4 \sum_{i=1}^{n} \Big((y_i - (x_i + \beta)^2) - 2(x_i + \beta)^2 \Big) = -4 \sum_{i=1}^{n} \big(y_i - 3(x_i + \beta)^2 \big)$.

We want to find a critical value of f, *i.e. find a root of* g, *i.e. find* β *such that* $g(\beta) = 0$, *using the Newton-Raphson method.*

The R program is as follows:

```
g = function(beta) { - 4 * sum( (y-(x+beta)∧2) * (x+beta) ) }
gp = function(beta) { - 4 * sum( y - 3 * (x+beta)∧2 ) }
# (Or, could omit "- 4" from both functions, since it cancels.)
betalist = c(2) # the initial value
for (i in 1:100) {
   betaprev = betalist[length(betalist)]
   betanew = betaprev - g(betaprev) / gp(betaprev)
   betalist = c(betalist, betanew)
}
print(betanew)
```

6. [10 points] Suppose we observe three (x, y) pairs: (1,4), (2,6), and (3,10). Compute (with full explanation) the cross-validation sum of squares (CVSS) for the unconstrained linear model, $Y = \beta_1 + \beta_2 X + \text{error}$ with β_1 and β_2 unknown. [Hint: remember that an unconstrained linear fit of two points passes through both points.]

Solution: $f_{-1}(x)$ *is an unconstrained linear fit of the two points (2,6) and (3,10), so* $f_{-1}(x)$ *passes through (2,6) and (3,10). Hence,* $f_{-1}(x) = 4x - 2$, *so* $f_{-1}(1) = 4(1) - 2 = 2$.

Similarly, $f_{-2}(x)$ *passes through (1,4) and (3,10), so* $f_{-2}(x) = 3x + 1$, *so* $f_{-2}(2) = 3(2) + 1 = 7$.

And, $f_{-3}(x)$ *passes through (1,4) and (2,6), so* $f_{-3}(x) = 2x + 2$, *so* $f_{-3}(3) = 2(3) + 2 = 8$.

Hence, $CVSS = \sum_{i=1}^{3} (y_i - f_{-i}(x_i))^2 = (4-2)^2 + (6-7)^2 + (10-8)^2 = 4 + 1 + 4 = 9$.

7. [10 points] Consider the estimator $\hat{\theta}$ which is "the second-largest value", i.e. $\hat{\theta}(x_1, \ldots, x_n) = \max\{x_i : x_i < \max(x_1, \ldots, x_n)\}$. [For example, $\hat{\theta}(3, 20, 17, 12, 15) = 17$.] Suppose we observe, from an unknown distribution, four data values: 5, 7, 8, 10. Compute (with explanation) $\widehat{\text{Var}}(\hat{\theta})$, the jackknife estimator of the variance of $\hat{\theta}$.

Solution: $\hat{\theta}_{-1} = \hat{\theta}(7, 8, 10) = 8$. *And,* $\hat{\theta}_{-2} = \hat{\theta}(5, 8, 10) = 8$. *And,* $\hat{\theta}_{-3} = \hat{\theta}(5, 7, 10) = 7$. *And,* $\hat{\theta}_{-4} = \hat{\theta}(5, 7, 8) = 7$. *So,* $\hat{\theta}_{\bullet} = \frac{1}{n} \sum_{i=1}^{n} \hat{\theta}_{-i} = \frac{1}{4}(8+8+7+7) = 7.5$.

Then, $\widehat{\text{Var}}(\hat{\theta}) = \frac{n-1}{n} \sum_{i=1}^{n} (\hat{\theta}_{-i} - \hat{\theta}_{\bullet})^2 = \frac{3}{4}\big((8-7.5)^2 + (8-7.5)^2 + (7-7.5)^2 + (7-7.5)^2\big) = \frac{3}{4}\big(0.25 + 0.25 + 0.25 + 0.25\big) = 3/4$.