STAT 428 – Homework #1
Luis S. Lin

Problem 1

```
> #### Problem 1
> #### Write R-code to calculate E(Nsd)
> #### for having ten sons and eight daughters
>
> # Initialize variables
> s=10; d=8; p=0.5; q=0.5
>
> # Create matrix and fill first column and row
> # Entry in matrix represents E[Nsd]
> x = matrix(nrow = s+1, ncol = d+1);
> x[1,] = c(1,(1:d)/q)
> x[,1] = c(1,(1:s)/p)
>
> # Iterate
> for (i in 1:s){
+ for (j in 1:d) {
+ x[i+1,j+1] = 1 + p*x[i,j+1] + q*x[i+1,j]
+ }
+ }
> # Display Results
> noquote("Expected number of children E(Nsd):"); print(x[s+1,d+1],
digits = 5)
[1] Expected number of children E(Nsd):
[1] 21.709
>
>
```

Problem 2

```
> #### Problem 2
> #### Write R function to compute
> #### S2n using recurrence relation
>
> # Find Variance using recursive expression for a given dataset
> # v = biased sample variance
> # m = sample mean
> # N = number of observations
> # x = dataset
> # n = iteration index
>
> findVar = function(x){
+
+ # Initialize variables
+ N = length(x);
+ m = x[1];
+ v = 0;
+
+ # Find variance recursively
+ for (n in 1:(N-1)){
+ # compute mean with n+1 observations
+ m = n/(n+1)*m + x[n+1]/(n+1);
+ # compute variance with n+1 observations
+ v = n/(n+1)*v + 1/n*(x[n+1] - m)^2;
+ }
+
+ # Output results
+ return(v)
+ }
>
> # Test function
>
> data = sample(1:100, 20, replace=T)
> findVar(data)
[1] 732.3
>
> sum((data-mean(data))^2)/length(data)
[1] 732.3
> # The results from the recursive method agree with that of the formula!
```

**** Note: S_n^2 recursive formulas in the book and lecture notes refer to the biased sample variance***

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_n)^2,$$

(the unbiased sample variance can be easily found from the biased one):

Problem 3

PROBLEM 3

$$E[N_{sd}^2] = \sum_{n=s+d}^{\infty} n^2 p(N_{sd}=n) =$$

$$= \sum_{n=s+d}^{\infty} n^2 \Big[P(N_{sd}=n \mid N_{s-1,d}=n-1) \times P(N_{s-1,d}=n-1) + P(N_{sd}=n \mid N_{s,d-1}=n-1) \times P(N_{s,d-1}=n-1) \Big]$$

$$= \sum_{n=s+d}^{\infty} n^2 p\, P(N_{s-1,d}=n-1) + \sum_{n=s+d}^{\infty} n^2 q\, P(N_{s,d-1}=n-1)$$

Let $m=n-1$

$$= p \sum_{m=s+d-1}^{\infty} (m+1)^2 P(N_{s-1,d}=m) + q \sum_{m=s+d-1}^{\infty} (m+1)^2 P(N_{d-1,s}=m)$$

$$= p\left\{ \sum_{m=s+d-1}^{\infty} m^2 P(N_{s-1,d}=m) + \sum_{m=s+d-1}^{\infty} 2m\, P(N_{s-1,d}=m) + \sum_{m=s+d-1}^{\infty} 1\cdot P(N_{s-1,d}=m) \right\}$$

$$+ q\left\{ \sum_{m=s+d-1}^{\infty} m^2 P(N_{s,d-1}=m) + \sum_{m=s+d-1}^{\infty} 2m\, P(N_{s,d-1}=m) + \sum_{m=s+d-1}^{\infty} 1\cdot P(N_{s,d-1}=m) \right\}$$

$$= p\left\{ E[N_{s-1,d}^2] + 2E[N_{s-1,d}] + 1 \right\} + q\left\{ E[N_{s,d-1}^2] + 2E[N_{s,d-1}] + 1 \right\}$$

$$\boxed{E[N_{sd}^2] = pE[N_{s-1,d}^2] + qE[N_{s,d-1}^2] + 2pE[N_{s-1,d}] + 2qE[N_{s,d-1}] + 1}$$

$s=0$: $N_{0d} \sim$ Negative Binomial (*) $\rightarrow E[N_{0d}] = d/q$, $Var[N_{0d}] = \frac{d-dq}{q^2}$

$Var[N_{0d}] = E[N_{0d}^2] - E[N_{0d}]^2 \rightarrow E[N_{0d}^2] = \frac{d-dq}{q^2} + \left(\frac{d}{q}\right)^2$

$d=0$: $N_{s0} \sim$ Negative Binomial (*) $\rightarrow$ $E[N_{s0}] = s/p$, $Var[N_{s0}] = \frac{s-sp}{p^2}$

$Var[N_{s0}] = E[N_{s0}^2] - E[N_{s0}]^2 \rightarrow E[N_{s0}^2] = \frac{s-sp}{p^2} + \left(\frac{s}{p}\right)^2$

∴ For both cases:

$$\boxed{E[N_{0d}^2] = \frac{d(1+d-q)}{q^2} \qquad E[N_{s0}^2] = \frac{s(1+s-p)}{p^2}}$$

(*) $N_{0d} \sim \binom{n-1}{d-1} q^{d-1} (1-q)^{n-d}$

(*) $N_{s0} \sim \binom{n-1}{s-1} p^{s-1} (1-p)^{n-s}$

Note: X = # trials for r successes

$P(X=x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$

$Y = X - r$ (# failures before r^{th} success)

$\Rightarrow P(Y=y) = \binom{r+y-1}{y} p^r (1-p)^y$

$E(Y) = \frac{r(1-p)}{p}$, $Var(Y) = \frac{r-rp}{p^2}$

$E(X) = E(Y) + E(r) = E(Y) + r = r/p$

$Var(X) = Var(Y) + \underbrace{\cancel{Var(r)}}_{0} = \frac{r-rp}{p^2}$

Problem 4

PROBLEM 4

Given: $\begin{cases} y_n = \int_0^1 \frac{x^n}{x+a}\,dx \\ y_n = \frac{1}{n} - a\,y_{n-1} \end{cases}$ $\rightarrow y_0 = \ln(x+a)\Big|_0^1 = \ln(1+a) - \ln(a)$

$\rightarrow y_0 = \ln\left(1 + \frac{1}{a}\right)$

PROVE: $y_n = \sum_{k=0}^{\infty} \frac{(-1)^k}{(n+k+1)\,a^{k+1}}$

Thus, y_0 case proven

PROOF by induction

1) $\underline{n=0}$: $y_0 = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)\,a^{k+1}} = \ln\left(1+\frac{1}{a}\right)$ $\left(a>1 \rightarrow 0<\frac{1}{a}<1\right)$

since $\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)}\,x^{k+1}$ for $|x|<1$.

2) Assume that case $\underline{i=n-1}$ (arbitrary case i) holds : $y_{n-1} = \sum_{t=0}^{\infty} \frac{(-1)^t}{(n+t)\,a^{t+1}}$

3) Show that case $\underline{i+1=n}$ holds:

Use given recursive relationship:

$$y_n = \frac{1}{n} - a\,y_{n-1} = \frac{1}{n} - a \sum_{t=0}^{\infty} \frac{(-1)^t}{(n+t)\,a^{t+1}} = \frac{1}{n} - \sum_{t=0}^{\infty} \frac{(-1)^t}{(n+t)\,a^t} =$$

$$= \cancel{\frac{1}{n}} - \cancel{\frac{1}{n}} - \sum_{t=1}^{\infty} \frac{(-1)^t}{(n+t)\,a^t} \underset{t=k+1}{=} \cancel{-} \sum_{k=0}^{\infty} \frac{(-1)^k \cancel{(-1)^1}}{(n+k+1)\,a^{k+1}}$$

$$\Rightarrow y_n = \sum_{k=0}^{\infty} \frac{(-1)^k}{(n+k+1)\,a^{k+1}}$$

$\therefore$ PROVEN by induction that $y_n = \sum_{k=0}^{\infty} \frac{(-1)^k}{(n+k+1)\,a^{k+1}}$ holds for all cases $(\forall n)$.