Homework # 6 *Statistics 134, Bandyopadhyay, Spring 2014*

3.1.12 a) If we think of N_i as just counting the number of times we get category i in n trials, we don't care what happens when we don't get this category. We get category i with probability p_i, and don't get it with probability $1 - p_i$, so the distribution of N_i is just binomial (n, p_i).

b) Similarly, $N_i + N_j$ counts the number of times we get either category i or j, which happens with probability $p_i + p_j$, so the distribution of $N_i + N_j$ is binomial $(n, p_i + p_j)$.

c) Now we just consider the three categories i,j, and everything else, which gives a joint distribution which is multinomial $(n, p_i, p_j, 1 - p_i - p_j)$.

3.1.24 a) Let $p_i = P(X = i \bmod 2)$, $i = 0, 1$. Then $p_0 + p_1 = 1$,

$$P(X + Y \text{ is even}) = p_0^2 + p_1^2 = 1 - 2p_0(1 - p_0) \geq 1 - 2 \cdot \frac{1}{4}.$$

b) Let $p_i = P(X = i \bmod 3)$, $i = 0, 1, 2$. Then $p_0 + p_1 + p_2 = 1$,

$$P(\text{X + Y + Z is a multiple of 3}) = p_0^3 + p_1^3 + p_2^3 + 6p_0p_1p_2.$$

Now write simply p, q, r for p_0, p_1, p_2. So

$$p + q + r = 1, 0 \leq p, q, r \leq 1.$$

To show: $p^3 + q^3 + r^3 + 6pqr \geq \frac{1}{4}$. Consider

$$1 = (p + q + r)^3 = p^3 + q^3 + r^3 + 6pqr + 3[p^2q + p^2r + q^2p + q^2r + r^2p + r^2q].$$

Notice that $p(q + r) = pq + pr \leq 1/4$, $q(p + r) = qp + qr \leq 1/4$, and $r(p + q) = rp + rq \leq 1/4$. The probability in question is thus

$$1 - 3[p[pq + pr] + q[qp + qr] + r[rp + rq]]$$
$$\geq 1 - 3[(p + q + r) \cdot 1/4] \geq 1 - 3/4 = 1/4.$$

3.2.14 We want $E(N)$, where N is the number of floors at which the elevator makes a stop to let out one or more of the people. N is a counting variable, so write $N = \sum_{j=1}^{10} I_j$ where I_j is the indicator of the event that at least one person gets off at floor j. By additivity,

$$E(N) = \sum_{j=1}^{10} E(I_j) = \sum_{j=1}^{10} P(\text{at least one person chooses floor } j).$$

Now for each j

$P(\text{at least one person chooses floor } j) = 1 - P(\text{nobody chooses floor } j) = 1 - (9/10)^{12}$

by the independence of the people's choices. Hence

$$E(N) = 10 \times \left[1 - (9/10)^{12}\right] \approx 7.18.$$

3.2.16 a) $P(X_i = k) = P(X_1 = k) = P(\text{first } k \text{ cards are non-aces, next is ace}) = \frac{(48)_k \cdot 4}{(52)_{k+1}}$.

b) $X_1 + X_2 + X_3 + X_4 + X_5 + 4 = 52 \Longrightarrow 5E(X_1) = 48 \Longrightarrow E(X_i) = E(X_1) = 9.6$.

c) No. For instance, $P(X_1 = 30, X_2 = 30) = 0$, but $P(X_1 = 30) = P(X_2 = 30) > 0$.

3.2.22 a) Let $I_{3,i}$ be the indicator of a run of length 3 starting at the ith trial. Let $P(3, i) = E(I_{3,i})$ be the probability of this event. Then for $n > 3$

$$\begin{aligned}
E(R_{3,n}) &= E(\sum_{i=1}^{n-2} I_{3,i}) = \sum_{i=1}^{n-2} E(I_{3,i}) = \sum_{i=1}^{n-2} P(3, i) \\
&= P(3, 1) + \sum_{i=2}^{n-3} P(3, i) + P(3, n-2) \\
&= p^3(1-p) + (n-4)(1-p)(p^3)(1-p) + (1-p)p^3 \\
&= 2p^3(1-p) + (n-4)(p^3)(1-p)^2
\end{aligned}$$

b) Similarly, let $I_{m,i}$ be the indicator of a run of length 3 starting at the ith trial. The for $m < n$

$$\begin{aligned}
E(R_{m,n}) &= \sum_{i=1}^{n-m+1} P(m, i) \\
&= 2p^m(1-p) + (n-m-1)(p^m)(1-p)^2
\end{aligned}$$

For $m = n$, $E(R_{n,n}) = p^n$.

c)

$$R_n = \sum_{m=1}^{n} R_{m,n}$$

so

$$\begin{aligned}
E(R_n) &= \sum_{m=1}^{n} E(R_{m,n}) \\
&= p^n + \sum_{m=1}^{n-1} 2p^m(1-p) + (n-m-1)(p^m)(1-p)^2 \\
&= p^n + 2(p-p^n) + (n-1)(1-p)^2 \sum_{m=1}^{n-1} (p^m) - (1-p)^2 \sum_{m=1}^{n-1} mp^m \\
&= 2p - p^n + \sum_{m=1}^{n-1} (n-m-1)(p^m)(1-p)^2
\end{aligned}$$

Let $\Sigma_1 = \sum_{m=1}^{n-1}(n-m-1)(p^m)(1-p)^2$. Then

$$\Sigma_1 - p\Sigma_1 = (n-1)p(1-p)^2 - \sum_{m=1}^{n-1} p^m(1-p)^2 = (n-1)p(1-p)^2 - (p-p^n)(1-p)$$

so

$$\Sigma_1 = (n-1)p(1-p) - (p-p^n)$$

and finally

$$E(R_n) = 2p - p^n + \Sigma_1 = p + (n-1)p(1-p)$$

d) Let I_j be the indicator that a run of some length starts on the jth trial. Then $R_n = \sum_{j=1}^{n} I_j$ and if we let $P(j)$ be the probability that a run of some length starts on the jth trial, we have

$$E(R_n) = E(\sum_{j=1}^{n} I_j) = \sum_{j=1}^{n} E(I_j) = \sum_{j=1}^{n} P(j)$$

where $P(1) = p$ and $P(j) = (1-p)p$ for $j > 1$. Thus

$$E(R_n) = p + (n-1)p(1-p)$$