Chap. 6, prob. 45. Order Statistics.
We have \( i.i.d. \) RV's \( X_1, X_2, X_3 \) uniformly distributed on \([0, 1]\). Let \( X_1 \leq X_2 \leq X_3 \) be the corresponding order statistics. We are asked to find \( P(X_1 + X_2 \leq X_3) \). We need the p.d.f. for the joint variables \((X_1, X_2, X_3)\), but this we can get from our previous work:

\[
f_{X_1, X_2, X_3}(x_1, x_2, x_3) = 3! \cdot 1, \text{ for } 0 \leq x_1 \leq x_2 \leq x_3 \leq 1, \text{ and 0 otherwise.}
\]

We have to integrate this p.d.f. over the “event” \( \{x_1 + x_2 \leq x_3\} \). The trick is describing the three variable integral correctly. We will integrate out the variable \( x_3 \) last, which runs from 0 to 1. With a fixed \( x_3 \in [0, 1] \), we have the following constraints on \( x_1, x_2 : x_1 \leq x_2 \leq x_3 - x_1 \) and \( x_1 \leq \frac{1}{2} x_3 \). (You should double check these! If you draw it – and you should! – you will get inside a square with side length \( x_3 \) one quarter of the square picked out by these inequalities. It will be the wedge from the two left side vertices joined to the center of the square.) Using this description gives:

\[
P(X_1 + X_2 \leq X_3) = \int_0^1 \int_0^{\frac{1}{2} x_3} \int_{x_1}^{x_3-x_1} 3! \cdot 1 \, dx_2 \, dx_1 \, dx_3
\]

\[
= 6 \int_0^1 \int_0^{\frac{1}{2} x_3} x_3 - 2x_1 \, dx_1 \, dx_3
\]

\[
= 6 \int_0^1 [x_3 x_1 - x_1^2]_0^{\frac{1}{2} x_3} \, dx_3
\]

\[
= 6 \int_0^1 \frac{1}{4} x_3^2 \, dx_3
\]

\[
= 6 \cdot \frac{1}{4} \cdot \frac{1}{3} = \frac{6}{12} = \frac{1}{2}.
\]

We have 100 people, with independent, equally likely days as their birthdays. The first part says: take 365 independent Bernoulli trials, one for each day,
where a success is when there are exactly three of our 100 people with their birthdays that day. Say $T_i, i = 1, \ldots, 365$, are the Bernoulli RV's for each of the 365 days. We want to know $E(T_1 + \ldots + T_{365}) = 365E(T_1)$, and $E(T_1) = \text{Prob(three births exactly on day 1)}$. This last is just

\[
\text{Prob(three births exactly on day 1)} = \binom{100}{3} \left(\frac{1}{365}\right)^3 \left(\frac{364}{365}\right)^{362}.
\]

Thus, the final answer is just $365 \cdot \binom{100}{3} \left(\frac{1}{365}\right)^3 \left(\frac{364}{365}\right)^{97}$.

**Chap. 7, prob. 46. Dice Game.**

First of all, we expect the two players’ fortunes to be correlated because they are connected via the bank rolling the dice to set the parameters of their game, similar to a problem we did in class. Basically, we just have to compute an expectation by conditioning. So, let $I_i, i = 1, 2$, be the RV's as in the problem, and let $B$ be the random variable “what the bank rolls”. We want to calculate $E(I_1) = E(I_2)$, and $E(I_1 \cdot I_2)$. So,

\[
E(I_1) = E(E(I_1|B = i)) = \sum_{i=2}^{12} \frac{12 - i}{11} \text{Prob}(B = i).
\]

Similarly,

\[
E(I_1 \cdot I_2) = E(E(I_1 \cdot I_2|B = i)) = \sum_{i=2}^{12} \left(\frac{12 - i}{11}\right)^2 \text{Prob}(B = i).
\]

Tune in later for the arithmetic! (Dec. 8, 2003, 9:45 AM.)

**Chap. 7, prob. 47. Random Graphs.**

The idea here is that you are given $n$ distinct points which can act as the vertices of a graph, that is, a configuration of vertices and edges joining the vertices. Our graphs will be random in the following way: we will say that at most one edge (say the straight line) will join a vertex $i$ to $j$. Whether this edge will be in the graph or not will be described by an indicator random variable $E_{i,j}$ which will be a Bernoulli RV with bias $p$. We will assume that $E_{i,j} = E_{j,i}$, so there is no effect of the direction of the edge (from $i$ to $j$, say, versus $j$ to $i$), and we assume that $E_{i,j}$ and $E_{i',j'}$ are independent if $(i,j) \neq (i',j')$ or $(j', i')$. Now the problem asks us to consider the “degree”
random variable $D_i$ for each vertex $i$, where

$$D_i = \sum_{j \neq i} E_{i,j},$$

that is, it counts the number of edges emanating from the $i$-th vertex. This gives us right away the answer to the first part of the question: each $D_i \sim \text{Binomial } [p, n - 1]$. To calculate the correlation coefficient, we should first think a bit intuitively about this. We can assume we are just talking about $D_1$ and $D_2$, since any pair will behave the same way. Now, if the random variables were independent, the correlation would be 0. These two are almost independent, since all but one of the possible edges for vertex 1 or 2 are independent. However, if the edge from 1 to 2 is an edge coming out of vertex 1, it is also an edge coming out of vertex 2, so we would expect the RV’s to be positively correlated. To see this, we first note that

$$E(D_1) = E(D_2) = (n - 1)p,$$

and

$$\text{Var}(D_1) = \text{Var}(D_2) = (n - 1)p(1 - p).$$

Write

$$D_1 = E_{1,2} + R_1,$$

where $R_1$ (the ”rest”) is the sum of the other edge indicator random variables for vertex 1. Similarly, $D_2 = E_{1,2} + R_2$, and $R_1$ and $R_2$ are independent. Now

$$\rho(D_1, D_2) = \frac{E(D_1 \cdot D_2) - E(D_1) \cdot E(D_2)}{\sqrt{\text{Var}(D_1)} \sqrt{\text{Var}(D_2)}}.$$

We know everything here except $E(D_1 \cdot D_2)$, but this is just given by

$$E(D_1 \cdot D_2) = E((E_{1,2} + R_1)(E_{1,2} + R_2)) = E(E_{1,2} + E_{1,2} \cdot R_1 + E_{1,2} \cdot R_2 + R_1 \cdot R_2)$$

$$= E(E_{1,2}) + E(E_{1,2}) \{E(R_1 + R_2)\} + E(R_1)E(R_2),$$

by the sum law and independence. To go further, observe that $R_1 \sim R_2 \sim \text{Binomial } [p, n - 2]$, so $E(R_1) = E(R_2) = (n - 2)p$. Putting this together, we get

$$E(D_1 \cdot D_2) = p + p \cdot \{2(n - 2)p\} + \{(n - 2)p\}^2 = (n - 1)^2p^2 + p(1 - p).$$
Finally, the correlation coefficient becomes
\[ \rho = \frac{(n - 1)^2 p^2 + p(1 - p) - (n - 1)^2 p^2}{(n - 1)p(1 - p)} = \frac{p(1 - p)}{(n - 1)p(1 - p)} = \frac{1}{n - 1}. \]

This is a very intuitive answer because it says that the two vertices are correlated exactly by the one common possible edge, which is fraction \( \frac{1}{n-1} \) of each vertex’s possible edges.

**Chap. 7, prob. 55. Ducks.**

We have 10 hunters, and an unspecified (random according to the Poisson\([6]\) distribution) number of ducks in a flock passing overhead. We have to figure what is the expected number of shot ducks, if each duck in the flock is chosen as target by the hunters independently of one another, with all ducks equally likely to be chosen by any hunter and with each hunter having a 0.6 probability of hitting a duck he or she aims at. (There is another, hidden, assumption we will clarify later.) This is another exercise in conditional expectation. We will use indicator variables to count the ducks: set \( D_i = 1 \) if the \( i \)-th duck is shot and 0 otherwise. Let \( N = \) the number of ducks in the flock, so that \( N \sim \text{Poisson}[6] \). As usual now, we set \( D = \sum_{i=1}^{N} D_i \) the RV “the number of ducks shot”. We want to calculate \( E(D) \). The first thing to notice is that it is **NOT** given by \( \sum_{i=1}^{N} E(D_i) \) (Why not?) This is where conditioning comes in:

\[
E(D) = E(E(D|N = n))
\]

\[
= E(NE(D_1|N = n))
\]

\[
= \sum_{n=0}^{\infty} nE(D_1|N = n)P(N = n)
\]

\[
= \sum_{n=0}^{\infty} nE(D_1|N = n)e^{-6} \frac{6^n}{n!}
\]

It remains to calculate \( E(D_1|N = n) \). We encounter the hidden assumption here: we will assume that if a hunter misses the duck he aims at, he will NOT hit another duck. So, the probability that somebody hits duck \#1 is \( 1 - \text{Prob(nobody hits it)} \). Now the probability that the first hunter doesn’t hit duck \#1 is the probability that he will not choose duck \#1, which is \( \frac{n-1}{n} \) together with the probability that the hunter chose duck \#1, but failed to hit it, which is \( \frac{0.4}{n} \). Altogether, this makes \( \frac{n-0.6}{n} \) the probability that the frist
hunter will not hit duck #1. So, altogether,

\[ E(D_1|N = n) = 1 - \left(\frac{n - 0.6}{n}\right)^{10}, \]

and so

\[ E(D) = \sum_{n=0}^{\infty} n(1 - \left(\frac{n - 0.6}{n}\right)^{10})e^{-6n} \frac{6^n}{n!}. \]

I don’t know much in the way of simplification for this, unfortunately.

Chap. 7, prob. 56. Elevators.
This is a little easier since we don’t have a so-called “random sum” as in #55. So there are \( N \) floors above the ground floor, and the number of riders \( R \) in the elevator entering at the ground level is random, but with a Poisson[10] distribution. Each of the passengers is equally likely to get out at any of the \( N \) possible floors, and this independently of what all the other passengers do. Compute the expected number of stops the elevator makes before discharging all its passengers. Note that this is going to be a lot like the birthdays problem above, except the number of people will be random. This makes it worth trying to compute the expectation by conditioning. First we introduce an indicator RV for each floor: \( F_i = 1 \) if somebody gets out at floor \( i \), and 0 otherwise. The number \( S \) of stops is given by \( S = \sum_{i=1}^{N} F_i \).

\[ E(S) = NE(F_1). \]

It is in computing \( E(F_1) \) that we will use conditioning.

\[ E(F_1) = E(E(F_1|R = r)). \]

However, \( E(F_1|R = r) \) is not so hard to compute:

\[ E(F_1|R = r) = \text{Prob}(F_1 = 1|R = r) = 1 - \left(\frac{N - 1}{N}\right)^r, \]
giving

\[ E(S) = NE(F_1) \]

\[ = \sum_{r=0}^{\infty} N \cdot \left\{ 1 - \left( \frac{N-1}{N} \right)^r \right\} \cdot e^{-10} \cdot \frac{10^r}{r!} \]

\[ = N \left\{ 1 - e^{-10} \cdot \sum_{r=0}^{\infty} \frac{(10(N-1))^r}{r!} \right\} \]

\[ = N \left\{ 1 - e^{-10} e^{10(N-1)/N} \right\} \]

\[ = N(1 - e^{-\frac{10}{N}}). \]