Recall: A geometrically distributed random variable X with constant probability of success p, $X \sim Geo(p)$ where we observe the trial number of the first success, $k = 1, 2, 3, \ldots$ has mean $E[X] = 1/p$ and variance $Var[X] = (1-p)/p^2$. (Note: different approach in text).

The geometric random variable is said to have no *memory*. To see what this means calculate the probability

$$Prob(Y > k) = 1 - Prob(Y \leq k)$$

Now calculate

$$Prob(Y > s + t \mid Y > s)$$

for positive integers s and t.

Example: Suppose that 20% of individuals in the Mall carry matches or a lighter.

- What is the probability that you will have to ask more than 3 people for a light till you get one?

- On average, how many people would you expect to ask?

- How many people should you plan on asking so that your probability of success is at least 0.75?

Poisson Random Variables

Adapted from an example in *Introductory Probability and Statistical Applications* by Paul L. Meyer.

Your company desires a presence on the Internet and so you have created a home page for your company. As you monitor hits on your homepage (occasions when someone accesses your information) you discover that, on average, you get 270 hits during a 3 hour period.

What is the probability that, during the next 3 minutes, you will receive 0 hits? 1 hit? 2 hits? ... 20 hits?...

First Attempt at a Solution
Since we are experienced with the binomial random variable, let's try to model this situation with n independent trials, each of which has a constant probability of success p. Divide the 3 minute interval into 9 subintervals of length 20 seconds each.

- How many hits would you expect in a 3 minute period?
- We are assuming that the probability of getting hit on any given 20 second period is the same. For each 20 second period, what is the probability of a success (you get a hit)?

What is you binomial model?

Using this model we obtain the probabilities

k, number of hits	p(k), probability
0	0.001953
1	0.017578
2	0.070312
3	0.164062
4	0.246094
5	0.246094
6	0.164062
7	0.070312
8	0.017578
9	0.001953

What are the limitations of this model? How may we improve the model?

Second Attempt at a Solution
Divide the 3 minute interval into 36 subintervals of length 5 seconds each.

Third Attempt at a Solution
Divide the 3 minute interval into 180 subintervals of length 1 second each.

Final Attempt at a Solution
Divide the 3 minute interval into many, say n, subintervals of length $(180/n)$ seconds each. Keep np constant, $np = 4.5$. For our special case one obtains:

In general, define the *intensity* as $\alpha = np$. Then, since we will compute a limit on binomial probabilities, we start with

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{1}$$

$$p(k) = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \tag{2}$$

$$p(k) = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} (\frac{\alpha}{n})^k (1-\frac{\alpha}{n})^{n-k} \tag{3}$$

$$p(k) = \frac{\alpha^k}{k!} \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} (1-\frac{\alpha}{n})^{n-k} \tag{4}$$

$$p(k) = \frac{\alpha^k}{k!} (\frac{n}{n})(\frac{n-1}{n})(\frac{n-2}{n})\dots(\frac{n-k+1}{n})(1-\frac{\alpha}{n})^n (1-\frac{\alpha}{n})^{-k} \tag{5}$$

$$p(k) = \frac{\alpha^k}{k!} [(\frac{n}{n})(\frac{n-1}{n})(\frac{n-2}{n})\dots(\frac{n-k+1}{n})(1-\frac{\alpha}{n})^{-k}](1-\frac{\alpha}{n})^n \tag{6}$$

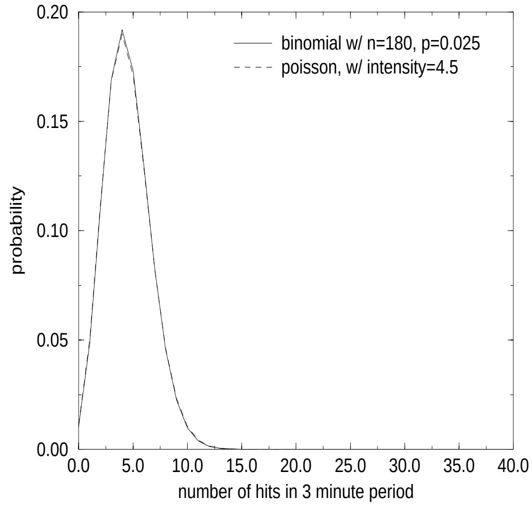
Take a limit on $p(k)$ as $n \to \infty$ and $np = \alpha = constant$. Note that, of course, $p = \alpha/n \to 0$. Each term in the square brackets approaches one. It is well known that

$$\lim_{n\to\infty} (1-\frac{\alpha}{n})^n \to e^{-\alpha}$$

Thus, in the limit,

$$prob(k\ hits) = \frac{e^{-\alpha}(\alpha)^k}{k!} = \frac{e^{-np}(np)^k}{k!}$$

The graph shows the binomial model with $n = 180$ and $p = 0.025$. On the same graph we see the Poisson model with $intensity = \alpha = 4.5$. The graphs are nearly identical.



NOTE: The Poisson random variable can assume values 0, 1, 2, 3, ..., i, ... There are an infinite number of possible values.

Examples: (1) A study of the effect on predator introduction to artificial ponds was published in the journal *Ecology.* Experimenters found that the average density of zoo-plankton in the pond was 4.60 individuals per centiliter. Assuming a Poisson distribution applies, what is the probability that a centiliter of fluid from the pond used in the study had

1. No individuals?
2. One individual?
3. Five individuals?
4. 150 individuals?

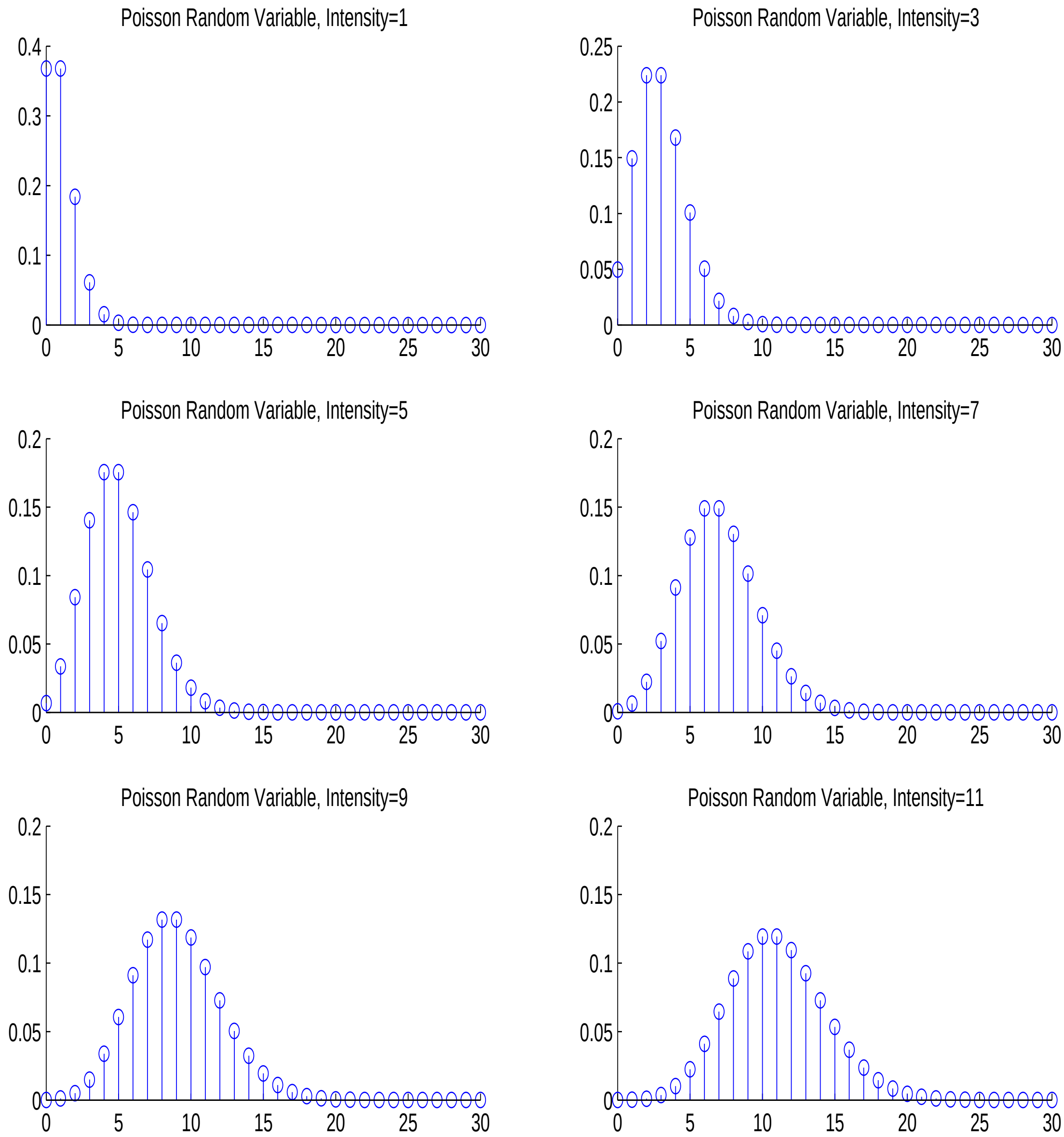
(2) Suppose a manufacturing process turns out items in such a way that a constant fraction of items, say p, are defective. If a lot of n such items is obtained, the probability of obtaining k defective items may be computed from the binomial formula. This is no problem if n and p are "manageable" numbers. We, however, have the following: Items are defective one time in a thousand. In a lot of 500 items, what is the probability of obtaining

1. No defective items?
2. One defective item?
3. Five defective items?
4. 150 defective items?

The following examples have been taken from Feller's *An Introduction to Probability Theory and its Applications*.

1. What is the probability that in a company of 500 employees exactly k will have birthdays on New Year's day?

2. Suppose that fasteners are manufactured under statistical quality control in such a way that it is legitimate to apply Bernoulli trials. Let the probability that a given fastener is defective be equal to $p = 0.015$. What is the probability that a box of 100 fasteners contains no defective items? (Evaluate this with a binomial model and with a Poisson model). How large should a box of fasteners be so that the probability that it contain at least 100 good fasteners is 0.8 or better? 0.9 or better? 0.95 or better?


Poisson Random Variable, Intensity=1
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
30
Poisson Random Variable, Intensity=3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Poisson Random Variable, Intensity=5
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Poisson Random Variable, Intensity=7
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Poisson Random Variable, Intensity=9
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Poisson Random Variable, Intensity=11
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30