Example: Suppose a discrete random variable is defined on the natural numbers 1, 2, ..., k, ... Is the probability distribution function defined as $p(k) = \left(\frac{1}{k}\right)^2$ a legitimate pdf? If not, how may we modify (normalize) it? What is it's expected value? (Recall reimann zeta function evaluated at one (harmonic series) and at 2 (our series)). What point are we trying to make?

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?

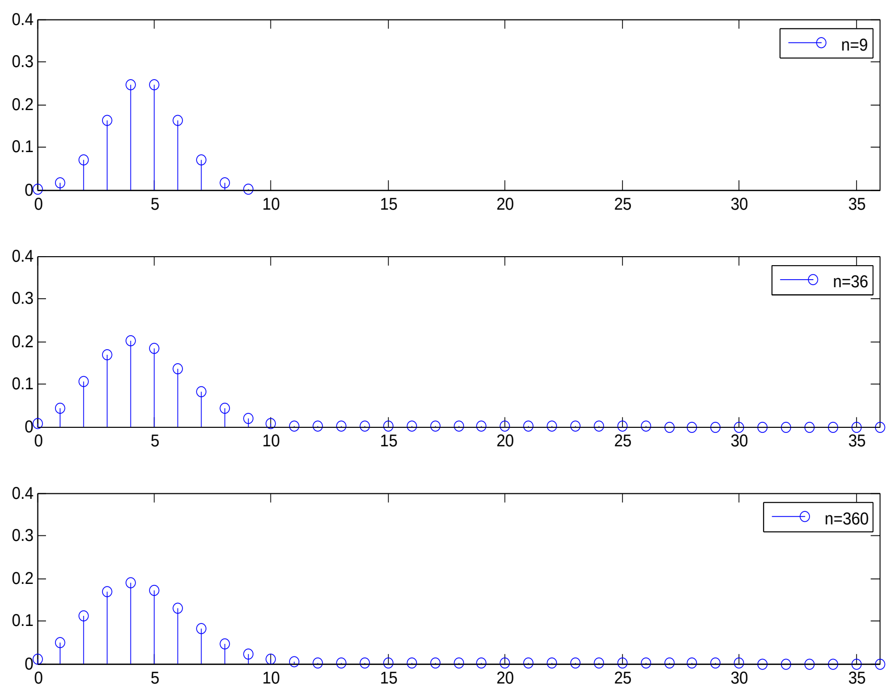
Using this model we obtain the probabilities

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

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 360 subintervals of length 0.5 seconds each.

The graph shows the binomial model with $n = 9$, $n = 36$, and $n = 360$, all with $E = 4.5$ and $p = E/n$. Note that there is some difference between the first two plots, but not very much between the second and third. We'll try to (analytically) find a limiting result.



Final Attempt at a Solution
Divide the 3 minute interval into many, say n, subintervals of length $(180/n)$ seconds each. Keep np constant, $np = 4.5$. Also, define the *intensity* as $\alpha = np$. Then, since we will compute a limit on binomial probabilities, we start with

$$
\begin{aligned}
p(k) &= \binom{n}{k} p^k (1-p)^{n-k} && (1)\\
p(k) &= \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} && (2)\\
p(k) &= \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}(\frac{\alpha}{n})^k(1-\frac{\alpha}{n})^{n-k} && (3)\\
p(k) &= \frac{\alpha^k}{k!}\frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}(1-\frac{\alpha}{n})^{n-k} && (4)\\
p(k) &= \frac{\alpha^k}{k!}(\frac{n}{n})(\frac{n-1}{n})(\frac{n-2}{n})\ldots(\frac{n-k+1}{n})(1-\frac{\alpha}{n})^n(1-\frac{\alpha}{n})^{-k} && (5)\\
p(k) &= \frac{\alpha^k}{k!}[(\frac{n}{n})(\frac{n-1}{n})(\frac{n-2}{n})\ldots(\frac{n-k+1}{n})(1-\frac{\alpha}{n})^{-k}](1-\frac{\alpha}{n})^n && (6)
\end{aligned}
$$

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,

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

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 No individuals? One individual? Five individuals? 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 no defective items? One defective item? Five defective items? 150 defective items?

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

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

(4) 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?