Lecture 2

Binomial and Poisson Probability Distributions

1) Bernoulli Distribution or Binomial Distribution:

Consider a situation where there are only two possible outcomes (these are called Bernoulli trials).

example: flipping a coin we either get a head or a tail
rolling a dice we either get a six or we do not get a six (we get 1, 2, 3, 4, 5)

Label the possible outcomes by the variable k. We want to find $P(k)$ the probability for event k to occur.
Since for this distribution k can take on only 2 values we define those values as:

$k = 0$ or $k = 1$

Let $P(k=0) = q$ (remember $0 \bullet q \bullet 1$), then we have:

$P(k=1) = p = 1 - q$ since $P(k=0) + P(k=1) = 1$ (*something* must happen)

We can write the probability distribution $P(k)$ as:

$P(k) = p^k q^{1-k}$ (this is a Bernoulli distribution)

Note: for the coin toss we could define $P(1)$ = probability for a head and assign its probability to be $P(1) = 0.5$
for the dice, define $P(1)$ = probability for a six to be rolled and assign its probability to be $P(1) = 1/6$

What is the mean (μ) of $P(k)$?

$$\mu = \frac{\sum_{k=0}^{1} P(k)k}{\sum_{k=0}^{1} P(k)} = \frac{0 \cdot q + 1 \cdot p}{q + p} = p$$

What is the Variance (σ^2) of $P(k)$?

$$\sigma^2 = \frac{\sum_{k=0}^{1} P(k)k^2}{\sum_{k=0}^{1} P(k)} - \mu^2 = (1-q) - (1-q)^2 = q(1-q) = qp$$

Suppose we have N trials (e.g. we flip a coin N times) what is the probability to get m successes (= heads)?

First consider tossing a coin twice. List all possible out comes from tossing to coins:

$m = 0$ (no heads) $P(m=0) = q^2$
$m = 1$ (one head) $P(m=1) = qp + pq$ (toss 1 is a tail, toss 2 is a head or toss 1 is head, toss 2 is a tail)
$m = 2$ (two heads) $P(m=2) = p^2$

Note that for the case where $m = 1$ we don't care which of the tosses is a head so there are <u>two</u> outcomes that give us 1 head. We can write $P(m=1)$ as:

$P(m=1) = 2pq$

We want the probability distribution function $P(m, N, p)$ where:

p = probability for a success (e.g. 0.5 for a head in a coin toss)
m = number of success (number of heads)
N = number of trials (e.g. coin tosses)

If we look at the three choices for the coin flip example, each term (m = 1, 2, 3) is of the form:

$C_m p^m q^{N-m}$ (m = 0, 1, 2), $N = 2$ for our example, $q = 1 - p$ always!

The coefficient C_m takes into account the number of ways an outcome can occur.
For m = 0 or 2 there is only one way for the outcome to arise (both tosses give heads or tails). Thus:

$C_0 = C_2 = 1$

For m = 1 (one head, two tosses) there are two ways that this can occur, thus: C_1 = 2.

These coefficients are called the "Binomial Coefficients"!
The Binomial Coefficients are just the number of ways of taking N things m at time and are given by:

$$C_{N,m} = \binom{N}{m} = \frac{N!}{m!(N-m)!} \qquad \text{Binomial Coefficient}$$

remember your factorials: 0! = 1! = 1, 2! = 1·2 = 2, 3! = 1·2·3 = 6,m! = 1·2·3··m

<u>Here the order of things is not important*.</u>
e.g. 2 tosses, one head case (m = 1), we don't care if toss 1 produced the head or if toss 2 produced the head.
* Unordered groups such as our example are often called *combinations*.

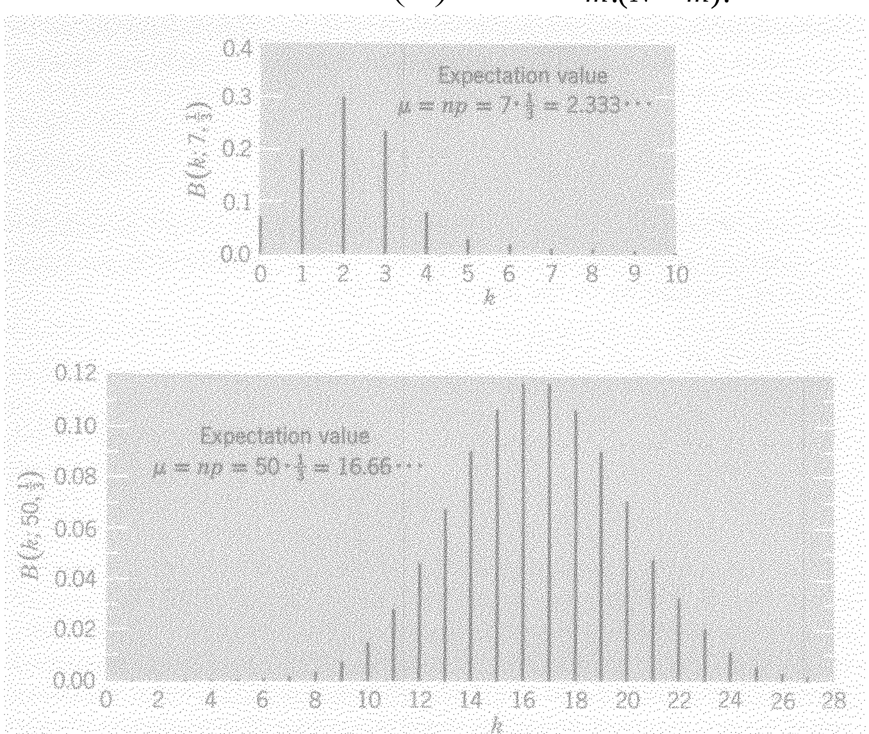
There are plenty of times where the order of occurrence is important, for example when we have N distinguishable objects and we want to group them m at a time. For example, if we tossed a coin twice (N = 2) and asked how many ways can we get one head (m = 1), the answer would be 2. Ordered arrangements are called *permutations* and are given by:

$$P_{N,m} = \frac{N!}{(N-m)!} \qquad \text{permutation}$$

<u>Example</u>: Suppose we have 3 balls, one white, one red, and one blue. The number of possible pairs we could have, keeping track of order is 6 (rw, wr, rb, br, wb, bw). Using the permutation formula we have: 3!/(3-2)! = 6. However, if order is <u>not</u> important (rw = wr) then we would use the binomial formula to get (3!)/[(3-2)!(2!)] = 3.

The binomial distribution is given by:

$$P(m,N,p) = C_{N,m}p^m q^{N-m} = \binom{N}{m}p^m q^{N-m} = \frac{N!}{m!(N-m)!}p^m q^{N-m}$$



Two binomial probability distributions. The asymmetry decreases as the expectation value, np, increases.

Does this formula make sense, e.g. if we sum over all possibilities do we get 1?
To show that this distribution is normalized properly, first remember the Binomial Theorem:

$$(a+b)^k = \sum_{l=0}^{k}\binom{k}{l}a^{k-l}b^l$$

For this example $a = q = 1 - p$ and $b = p$, and (by definition) $a + b = 1$.

$$\sum_{m=0}^{N} P(m,N,p) = \sum_{m=0}^{N} \binom{N}{m} p^m q^{N-m} = (p+q)^N = 1$$

Thus the distribution is normalized properly.

What is the mean of this distribution?

$$\mu = \frac{\sum_{m=0}^{N} mP(m,N,p)}{\sum_{m=0}^{N} P(m,N,p)} = \sum_{m=0}^{N} mP(m,N,p) = \sum_{m=0}^{N} m\binom{N}{m} p^m q^{N-m}$$

A cute way of evaluating the above sum is to take the derivative:

$$\frac{\partial}{\partial p}\left[\sum_{m=0}^{N} \binom{N}{m} p^m q^{N-m}\right] = 0 = \sum_{m=0}^{N} m\binom{N}{m} p^{m-1} q^{N-m} - \sum_{m=0}^{N} \binom{N}{m} p^m (N-m)(1-p)^{N-m-1}$$

$$\sum_{m=0}^{N} m\binom{N}{m} p^{m-1} q^{N-m} = \sum_{m=0}^{N} \binom{N}{m} p^m (N-m)(1-p)^{N-m-1}$$

$$p^{-1}\sum_{m=0}^{N} m\binom{N}{m} p^m q^{N-m} = (1-p)^{-1} N \sum_{m=0}^{N} \binom{N}{m} p^m (1-p)^{N-m} - (1-p)^{-1} \sum_{m=0}^{N} \binom{N}{m} m p^m (1-p)^{N-m}$$

$$p^{-1}\mu = (1-p)^{-1} N(1) - (1-p)^{-1}\mu$$

$\mu = Np$ for a binomial distribution

What's the variance of a binomial distribution?

Using a similar trick as for the average we find:

$$\sigma^2 = \frac{\sum_{m=0}^{N} (m-\mu)^2 P(m,N,p)}{\sum_{m=0}^{N} P(m,N,p)} = Npq$$

Example 1: Suppose you observed m special events in a sample of N events. The measured probability (sometimes called "efficiency") for a special event to occur is $\varepsilon = m/N$. The error on the probability is

$$\delta\varepsilon = \frac{\sqrt{N\varepsilon(1-\varepsilon)}}{N} = \sqrt{\frac{\varepsilon(1-\varepsilon)}{N}}$$ (sometimes called "error on the efficiency")

Thus you want to have a sample (N) as large as possible to reduce the certainty in the probability measurement!

Example 2: Suppose a baseball player's batting average is 0.333 (1 for 3 on average).
Consider the case where the player either gets a hit or makes an out (forget about walks here!).
In this example: p = prob. for a hit = 0.333 and $q = 1 - p = 0.667$ (prob. for "no hit").
On average how many hits does the player get in 100 at bats?

$\mu = Np = 100(0.33) = 33$ hits

What's the standard deviation for the number of hits in 100 at bats?

$\sigma = (Npq)^{1/2} = (100\cdot0.33\cdot0.67)^{1/2}$ • 4.7 hits

Thus we expect • 33 ± 5 hits per 100 at bats
Consider a game where the player bats 4 times:

Probability of 0/4 = $(0.67)^4$ = 20%
Probability of 1/4 = $[4!/(3!1!)](0.33)^1(0.67)^3$ = 40%
Probability of 2/4 = $[4!/(2!2!)](0.33)^2(0.67)^2$ = 29%

Probability of 3/4 = [4!/(1!3!)](0.33)^3(0.67)^1 = 10%
Probability of 4/4 = [4!/(0!4!)](0.33)^4(0.67)^0 = 1%

Note: the probability of getting at least one hit is: 1 - P(0) = 0.8

2) The Poisson Distribution:

Another popular discrete distribution is the Poisson distribution. Consider the following conditions:

a) p is very small and approaches 0. For example suppose we had a 100 sided dice instead of a 6 sided dice. Here p = 1/100 instead of 1/6. Suppose we had a 1000 sided dice, p = 1/1000...etc.

b) N is very large, it approaches •. For example, instead of throwing 2 dice, we could throw 100 or 1000 dice.

c) The product Np is finite.

A good example of the above conditions occurs when one considers radioactive decay.

Suppose we have 25 mg of an element. This is • 10^{20} atoms.

Suppose the lifetime (λ) of this element = 10^{12} years • 5x10^{19} seconds.

The probability of a given nucleus to decay in one second = 1/λ = 2x10^{-20}/sec.

For this example: $N = 10^{20}$ (very large)
$p = 2\text{x}10^{-20}$ (very small)
$Np = 2$ (finite!)

We can derive an expression for the Poisson distribution by taking the appropriate limits of the binomial distribution.

$$P(m,N,p) = \frac{N!}{m!(N-m)!}p^m q^{N-m}$$

Using condition b) we obtain:

$$\frac{N!}{(N-m)!} = \frac{N(N-1)\cdots(N-m+1)(N-m)!}{(N-m)!} \approx N^m$$

$$q^{N-m} = (1-p)^{N-m} = 1 - p(N-m) + \frac{p^2(N-m)(N-m-1)}{2!} + \ldots \approx 1 - pN + \frac{(pN)^2}{2!} - \approx e^{-pN}$$

Putting this altogether we obtain:

$$P(m,N,p) = \frac{N^m p^m e^{-pN}}{m!} = \frac{e^{-\mu}\mu^m}{m!}$$

Here we've let $\mu = pN$.

It is easy to show that:

$\mu = Np$ = mean of a Poisson distribution

$\sigma^2 = Np = \mu$ = variance of a Poisson distribution.

Note: m is always an integer • 0.

μ does *not* have to be an integer.

Since Np is fixed, we write $P(m, \mu)$ instead of $P(m, N, p)$ (only need 2 numbers to specify Poisson)

Back to our example of radioactivity.

a) What's the probability of zero decays in one second if the average = 2 decays/sec?

$$P(0,2) = \frac{e^{-2}2^0}{0!} = \frac{e^{-2}1}{1} = e^{-2} = 0.135 \rightarrow 13.5\%$$

b) What's the probability of more than one decay in one second if the average = 2 decays/sec?

$$P(>1,2) = 1 - P(0,2) - P(1,2) = 1 - \frac{e^{-2}2^0}{0!} - \frac{e^{-2}2^1}{1!} = 1 - e^{-2} - 2e^{-2} = 0.594 \rightarrow 59.4\%$$

c) Estimate the most probable number of decays/sec?

We want: $\frac{\partial}{\partial m}P(m,\mu)\Big|_{m*} = 0$

To solve this problem its convenient to maximize ln$P(m, \mu)$ instead of $P(m, \mu)$.

$$\ln P(m,\mu) = \ln(\frac{e^{-\mu}\mu^m}{m!}) = -\mu + m\ln\mu - \ln m!$$

In order to handle the factorial when take the derivative we use *Stirling's Approximation*:

$$\ln(m!) \bullet m\ln(m)\text{-}m$$

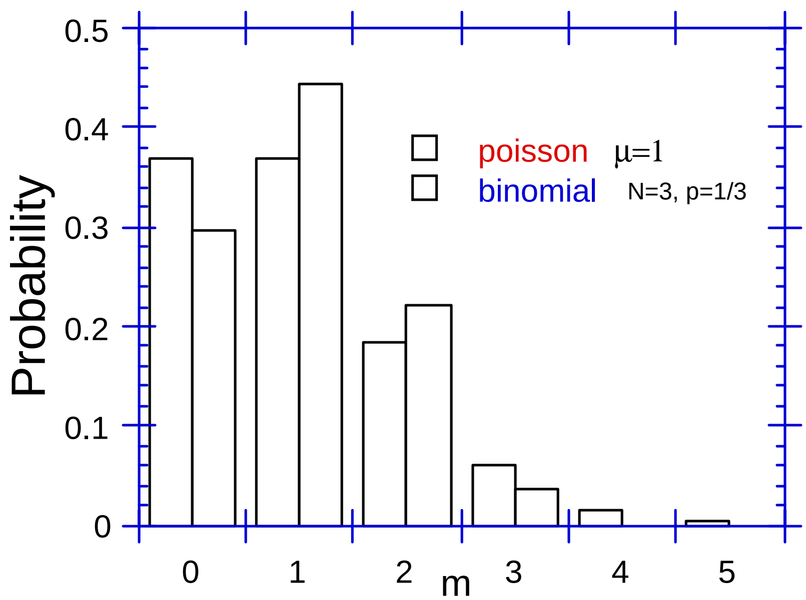
$$\frac{\partial}{\partial m}\ln P(m,\mu) = \frac{\partial}{\partial m}(-\mu + m^*\ln\mu - \ln m^*!) \approx \frac{\partial}{\partial m}(-\mu + m^*\ln\mu - m^*\ln m^* + m^*) = \ln\mu - \ln m^* - 1 + 1 = 0$$

$$m^* = \mu$$

In this example the most probable value for m is just the average of the distribution. <u>Therefore if you observed m events in an experiment, the error on m is $\sqrt{\mu} = \sqrt{m}$</u> .

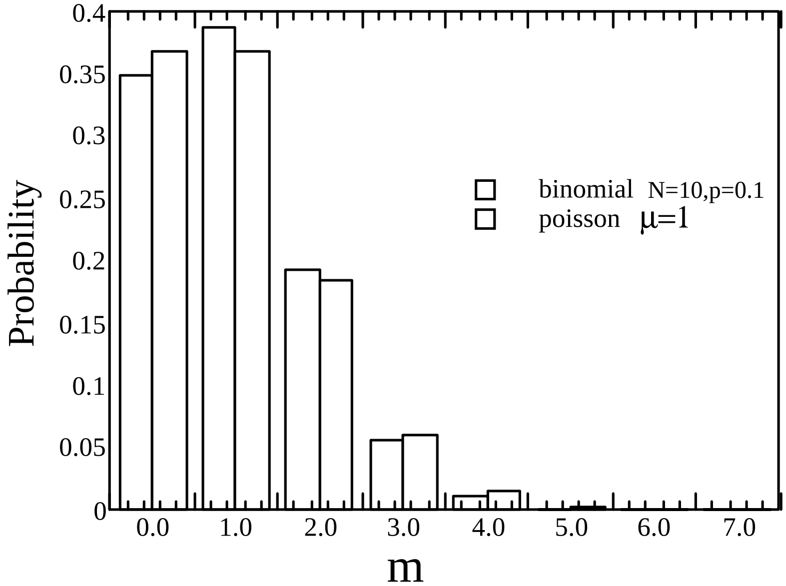
<u>Caution</u>: The above derivation is only approximate since we used Stirlings Approximation which is only valid for large m. Another subtle point is that strictly speaking m can only take on integer values while μ is not restricted to be an integer.

Below is a comparison of a binomial and Poisson distribution, each with a mean of 1.



Below is a comparison between a binomial and Poisson distribution where again the mean of each distribution =1.

As we see, for moderately small p (= 1/10) there's not much difference between the two distributions.


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Probability
binomial N=10,p=0.1
poisson μ=1
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
m