Probability reminders

BIOL 434/509

Mean and expected value

The expected value of a quantity is the mean of that value.

$$E[x] = Mean[x]$$

The mean is defined as a property of a probability distribution:

$\mu = E[x] = \sum x\, f(x)$, where $f(x)$ is the probability of x.

$\mu = E[x] = \int x\, \phi(x) dx$, where $\phi(x)$ is the probability density function of x.

Variance

$$var(x) = E[(x - \mu)^2] = \sum (x - \mu)^2 f(x)$$

$$var(x) = E[x^2] - \mu^2$$

The standard deviation is the positive square root of the variance:

$$sd(x) = \sqrt{var(x)}$$

Covariance

Covariance is a measure of association between two variables, related to variance.

$$cov(x, y) = E[(x - \mu_X)(y - \mu_Y)]$$

The covariance of a variable with itself is its variance.

Conditional probability

Conditional probability: the probability of one event *given that* some other information is true.

$\Pr[X \mid Y]$ means the probability of X given that Y is true.

$$\Pr[\text{A}\ allele\ in\ gamete \mid parent\ is\ \text{Aa}] = 1/2$$

Law of total probability

$$\Pr[X] = \sum \Pr[X \mid Y]\Pr[Y]$$

e.g.,

$$\begin{aligned}\Pr[\textit{gamete contains allele}\ \mathrm{A}] \quad = \\ &\Pr[A \mid \textit{parent is}\ \mathrm{AA}]\,\Pr[\textit{parent is}\ \mathrm{AA}] \\ &+ \Pr[A \mid \textit{parent is}\ \mathrm{Aa}]\,\Pr[\textit{parent is}\ \mathrm{Aa}] \\ &+ \Pr[A \mid \textit{parent is}\ \mathrm{aa}]\,\Pr[\textit{parent is}\ \mathrm{aa}]\end{aligned}$$

$$= (1 \times P_{AA}) + \left(\frac{1}{2} \times P_{Aa}\right) + (0 \times P_{aa}) = P_{AA} + \frac{P_{Aa}}{2}$$

Independence

If X and Y are independent, then

$$P(X \text{ and } Y) = P(X)\,P(Y)$$

Binomial distribution

n independent trials;

for each trial there is a probability p of success;

Binomial distribution gives distribution of the total number of successes X

Binomial distribution

$$f(X) = \binom{n}{X} p^X (1-p)^{n-X},$$

where $\binom{n}{X} = n!/(X!\,(n-X)!)$ is the binomial coefficient, the number of orders that *X* successes can appear out of *n* trials.

$$X! \;=\; X\times(X-1)\times(X-2)\ldots\times2\times1$$

Note that 1! =1   and   0!=1

Binomial distribution

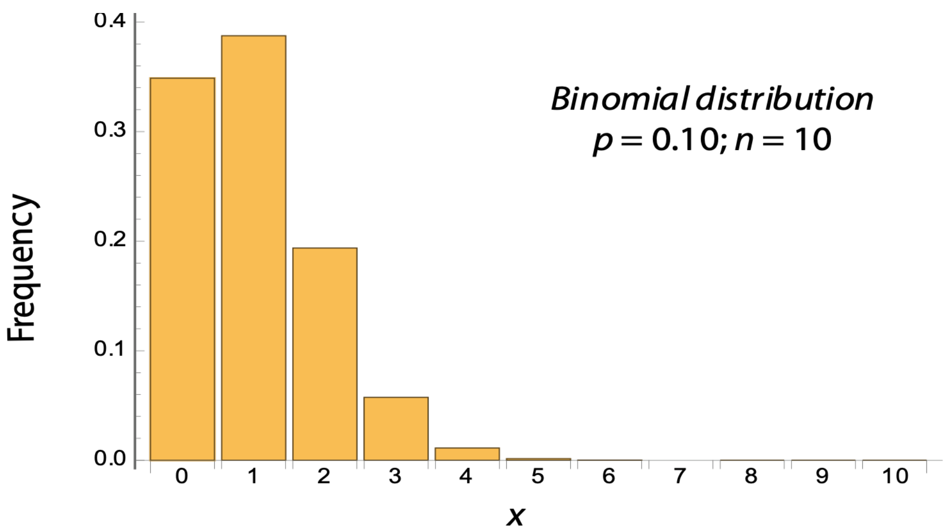
Mean number of successes:

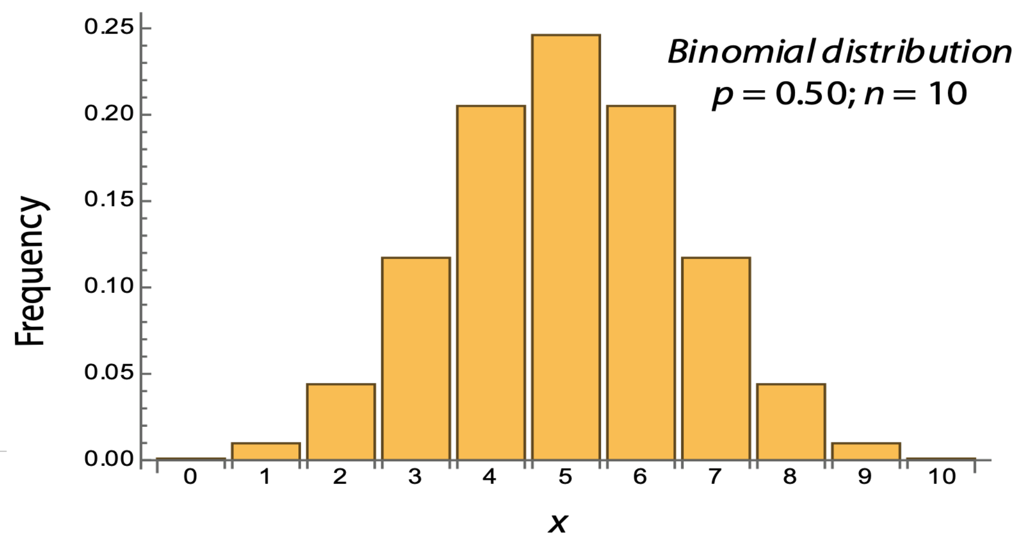
$$\mathrm{E}[X] = np.$$

Variance of the number of successes:

$$\mathrm{var}[X] = np(1-p).$$

Binomial distribution





Poisson distribution

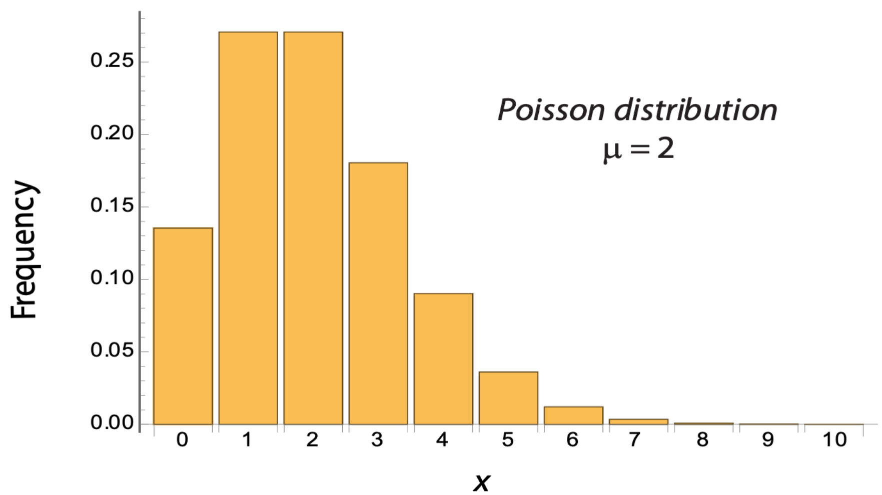
The **Poisson distribution** described the number of events that happen per unit of time or space, assuming all events are independent and happen with equal likelihood in all units.

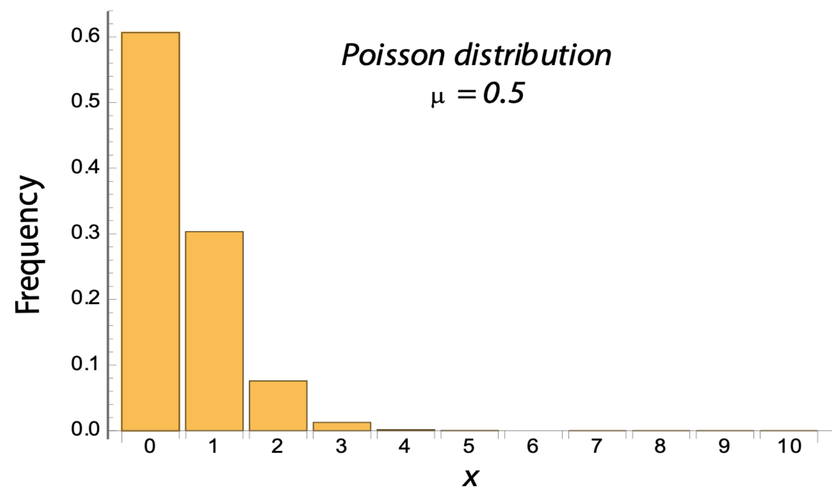
Poisson distribution

$$f(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

The mean and variance of the Poisson distribution are both μ.

Poisson distribution





Geometric distribution

If we do a series of independent trials, and for each trial there is a probability of success *p*, the **number of trials *t* until the first success** follows a geometric distribution.

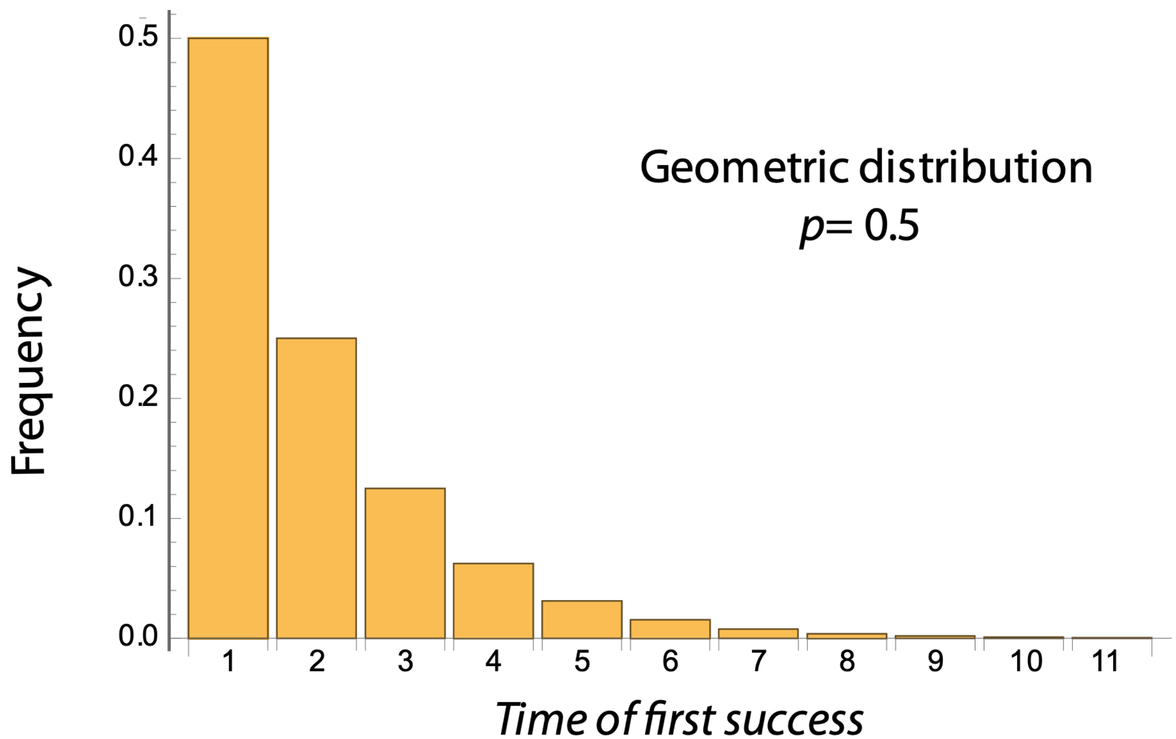
The probability that the first success occurs in the t^{th} trial is:

$$f(t) = (1-p)^{t-1}p$$

Mean: $1/p$

Variance: $\frac{1-p}{p^2}$

Geometric distribution



Useful equations for mean and variance

	Mean	Variance
$X + Y$	$E[X] + E[Y]$	$var[X] + var[Y] + 2cov[X, Y]$
$-X$	$-E[X]$	$var[X]$
$a X$	$a E[X]$	$a^2 var[X]$
$X + a$	$E[X] + a$	$var[X]$