Binomial coefficients

- This is a generalization of the definition of binomial coefficients you are familiar with.
- **Definition:** For $x \in \mathbb{R}_+$ and $k \in \mathbb{N}_0$,

$$\binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k!}$$

- If $x > k - 1$, then

$$\binom{x}{k} = \frac{\Gamma(x+1)}{\Gamma(k+1)\Gamma(x-k+1)}$$

- In particular, for $n > k, n, k \in \mathbb{N}_0$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

A modeling problem

- Imagine a sequence of independent Bernoulli trials, such that:

i. each trial results in success or failure;
ii. the probability of success for each trial, p, is constant across the trials;
iii. the experiment continues until a fixed number of successes r has been achieved.

- Then, the total number of failures before the r^{th} success, called N, is recorded .

The negative binomial distribution

- **Definition:** Let the random variable N have the probability mass function

$$p_N(k) = \mathbb{P}[N = k] = \binom{k + r - 1}{k} \left(\frac{1}{1 + \beta}\right)^r \left(\frac{\beta}{1 + \beta}\right)^k \qquad k = 0, 1, 2, \ldots$$

  for some constants $r > 0, \beta > 0$. Then, we say that N has the negative binomial distribution with parameters β and r.
- The r.v. N exactly models the situation in our "modeling problem" with probability of success $p = 1/(1 + \beta)$.

The negative binomial distribution: pg.f., $\mathbb{E}$, Var

- The probability generating function of $N \sim NegBin(\beta, r)$ is

$$P_N(z) = [1 - \beta(z-1)]^{-r}$$

- The expected value is

$$\mathbb{E}[N] = r\beta$$

- The variance is

$$Var[N] = r\beta(1+\beta)$$

The geometric distribution

- **Definition:** The geometric distribution is a special case of the negative binomial distribution for $r = 1$.
- For $N \sim \textit{Geometric}(\beta)$, we have

$$p_N(k) = \mathbb{P}[N = k] = \frac{\beta^k}{(1+\beta)^{k+1}} \qquad k = 0, 1, 2, \ldots$$

- In words $p_N(k)$ is the probability of the first success happening in the $(k+1)^{st}$ trial

The geometric distribution (cont'd)

- Note: For $n \geq 0$,

$$\mathbb{P}[N > n] = \sum_{k=n+1}^{\infty} p_N(k)$$

$$= \frac{1}{1+\beta} \sum_{k=n+1}^{\infty} \left(\frac{\beta}{1+\beta}\right)^k = \left(\frac{\beta}{1+\beta}\right)^{n+1}$$

- The **geometric distribution** has the memoryless property (this is the property that - as we learned in Probability - the exponential distribution also has), i.e., for $m, n \geq 0$,

$$\mathbb{P}[N > m+n | N > m] = \mathbb{P}[N > n]$$