3.5 The geometric probability distribution

- Same setup as binomial experiments
- Y=(# of trials on which the first success occurs)
- Possible values of Y?
- E_1: S,

  E_2: FS,

  E_3: FFS,

  …

  E_k: FFF…FFFS (k F's)

  …
- *$p(y)=P(Y=y)=P(E_y)=P(FFF\ldots FFFS)=qqq\ldots qqqp = q^{y-1}p$*
- Definition 3.8 A random variable y is said to have a geometric probability distribution iff

  $p(y)=q^{k-1}p$, $y=1,2,3,\ldots$ $0\le p\le 1$.
- Do they add up to 1? (Exercise 2.50)
- Probability histogram (use R; barplot(dgeom(1:7,p=.5))
- Often used to model distributions of lengths of waiting times.
- Example 3.11 The probability of engine malfunction during any 1-hour period is p=.02. What's P(a given engine will survive 2 hours)=?
  - Let Y=(# of 1-hour intervals until the first malfunction)
  - P(survive 2 hours) = $P(Y\ge 2)$
- *Theorem 3.8. If Y is a random variable with a geometric distribution,*

  $\mu=E(Y)=1/p$ and $\sigma^2=V(Y) = (1-p)/p2$
  - Proof:
- What's E(Y) and var(Y) for the Y above?
- HW. Some of the exercises 3.50~71
- Keywords: geometric distribution; Y~geom(p)

3.6 The negative binomial probability distribution

- Y=the number of the trial on which the r'th success occurs
- *Y~NB(r,p) has a probability function*

  $$p(y)=\binom{y-1}{r-1}p^r q^{y-r}, y=r,r+1,r+2,\ldots$$

  Also,

  $\mu=E(Y)=r/p$ and $\sigma^2=V(Y) = r(1-p)/p2$

- In R, run barplot(dnbinom(1:10, 5, .5))
- Skip in the current class but it's a distribution useful in many applications
- HW. Some of the exercises 3.72~83
- Keywords: negative binomial distribution; Y~nbinom(r,p)

3.7 The hypergeometric probability distribution

- Similar to binomial experiment but sampling *without* replacement. (cf. voter poll example above) Need to consider when n is large relative to N.
- Definition 3.10 A random variable Y is said to have a hypergeometric probability distribution if and only if dist $p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$ where $y = 0,1,\ldots,n$ subject to y≤r and n-y ≤ N-r.
- *Derivation?*
- Example 3.16 Select 10 engineers randomly from a group of 20. What's P(the 10 selected include all the 5 best engineers)
- *Theorem 3.10 For Y~hypergeometric,* $\mu = E(Y) = \frac{nr}{N}$ *and* $\sigma^2 = V(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$.
- Resemblance to binomial distribution can be seen by letting p=r/N and letting N->∞.
- Example 3.17 Lots of 20. Reject a lot if more than 1 defective is observed out of 5 samples from a lot. If a lot contains 4 defectives, what's the probability that it will be rejected? What are the E and V of the number of defectives in the sample of 5?
- HW. Some of the exercises 3.84~96
- Keywords: hypergeometric distribution; Y~hyper(# of 'success' balls, # of 'failure' balls, sample size)