3.5 **The geometric probability distribution**

- Same setup as binomial experiments
- Y=(# of trials on which the first success occurs)
- Possible values of Y?
  - E₁: S,
  - E₂: FS,
  - E₃: FFS,
  - ...
  - Eₖ: FFF…FFFS (k F’s)
- \( p(y)=P(Y=y)=P(E_y)=P(FFF…FFFS)=qqq…qqqp = q^{y-1}p \)
- Definition 3.8 A random variable y is said to have a geometric probability distribution iff \( p(y)=q^{y-1}p, y=1,2,3,… \) \( 0 \leq p \leq 1 \).
- 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 \( P(\text{a given engine will survive 2 hours})=? \)
  - Let Y=(# of 1-hour intervals until the first malfunction)
  - \( P(\text{survive 2 hours}) = P(Y \geq 2) \)
- **Theorem 3.8. If Y is a random variable with a geometric distribution,**
  - \( \mu=E(Y)=1/p \) and \( \sigma^2=V(Y) = (1-p)/p^2 \)
  - 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
- \( Y \sim NB(r,p) \) has a probability function
  
  \[
  p(y) = \binom{y-1}{r-1}p^r q^{y-r}, y=r,r+1,r+2,...
  \]
  
  Also,
  
  \( \mu=E(Y)=r/p \) and \( \sigma^2=V(Y) = r(1-p)/p^2 \)
1. In R, run `barplot(dnbinom(1:10, 5, .5))`
2. Skip in the current class but it’s a distribution useful in many applications
3. HW. Some of the exercises 3.72~83
4. 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

\[
P(y) = \frac{{\binom{r}{y} \binom{N-r}{n-y}}}{{\binom{N}{n}}} \quad \text{where} \quad y = 0, 1, \ldots, n \quad \text{subject to} \quad y \leq r \quad \text{and} \quad n-y \leq N-r.
\]

- **Derivation?**

- **Example 3.16** Select 10 engineers randomly from a group of 20. What’s \( P(\text{the 10 selected include all the 5 best engineers}) \)

- **Theorem 3.10** For \( Y \sim \text{hypergeometric} \), \( \mu = E(Y) = \frac{nr}{N} \) 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 \( N \to \infty \).

- **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; \( Y \sim \text{hyper}(\text{# of ‘success’ balls, # of ‘failure’ balls, sample size}) \)