2.4 The Hypergeometric & Multinomial Distributions

Hypergeometric Distribution

Out of a population of size N, suppose we have N_1 successes and N_2 failures.

(note, $N_1 + N_2 = N$, the probability of a success, $p = N_1 / N$)

Define a random variable X:

the number of successes in a random sample of size n.

If sampling is done without replacement, X follows a **hypergeometric distribution.**

Hypergeometric Distribution

$$X \sim Hypergeom(N, N_1, n)$$

(remember: $N = N_1 + N_2$*)*

$$f(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}},$$

$$x \leq n, \quad x \leq N_1, \quad n - x \leq N_2$$

- $E[X] = n\frac{N_1}{N}$
- $Var[X] = n\frac{N_1}{N}\frac{N_2}{N}\frac{N-n}{N-1}$

Hypergeometric vs Binomial

If instead, sampling is done one at a time with replacement, X $\sim$ Binomial(n,p) e.g.

- **<u>Binomial</u>:** A magical beer machine gives the user a stout 30% of the time and an IPA 70% of the time. Let X be the number of stouts you get out of 20 beers.
- **<u>Hypergeometric</u>:** A nice minibar has 9 stouts and 21 IPAs. Let X be the number of stouts you get if you randomly select 20 beers.
  - What is the pmf of X? $f(x) = \frac{\binom{9}{x}\binom{21}{5-x}}{\binom{30}{20}}, \;\; x \leq 9$

Multinomial Distribution

Similar to binomial distribution, but for more than 2 groups. E.g.

- Color – Red/Green/Blue
- Your Major – Stats/Math/Engineering/Other

Multinomial Distribution

$$X = (X_1, X_2, \dots X_k) \sim Multinomial(n, p_1, p_2, \dots, pk)$$

$$f(x_1, x_2, \dots x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1}\, p_2^{x_2} \dots p_k^{x_k}$$

- $E[X_i] = np_i$
- $Var[X_i] = np_i(1\text{-}p_i)$

Examples

2.4

1 A deck of cards contains 20 cards: 6 red cards and 14 black cards. 5 cards are drawn randomly *without replacement*.

What is the probability that exactly 4 red cards are drawn?

What is the probability that at least 2 black cards are drawn?

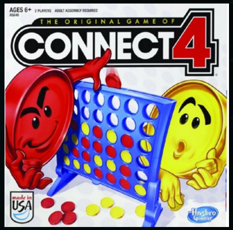
2

Suppose the majors of students taking Stat 400 can be broken down as follows:

Math	Statistics	Other
10%	20%	70%

Out of 10 randomly sampled students, calculate the probability that this group contains:

- A) 2 Math, 2 Stats, and 6 Other
- B) At least one Stats student

3 When Iron Man and Captain America play 

against each other, Iron Man wins 40% of the time, loses 35% of the time and draws 25% of the time. Assume results of games are independent.

If they play 12 games, what is the probability that Iron Man wins 7, loses 2, and draws 3 games?

$$f(x_1, x_2, \ldots x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1}\, p_2^{x_2} \ldots p_k^{x_k} = \frac{12!}{7!2!3!} \times .40^7 \times .35^2 \times .25^3 = 0.0248$$

If they play 12 games, what is the expected value of the number of games that they will tie?

2.6 The Poisson Distribution

Poisson Process

Definition 2.6-1

Let the number of occurrences of some event in a given continuous interval be counted. Then we have an **approximate Poisson process** with parameter $\lambda > 0$ if the following conditions are satisfied:

(a) The numbers of occurrences in nonoverlapping subintervals are independent.

(b) The probability of exactly one occurrence in a sufficiently short subinterval of length h is approximately λh.

(c) The probability of two or more occurrences in a sufficiently short subinterval is essentially zero.

Poisson Process Examples

- # of cell phone calls passing through a relay tower between 9 and 11 a.m.
- Number of customers that show up to Oberweis between 5-6pm.
- Number of defects in a 100-foot roll of aluminum screen that is 2 feet wide.

Poisson Distribution

$$X \sim Poisson(\lambda)$$

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0,1,2,\ldots$$

- $E[X] = \lambda$
- $Var[X] = \lambda$

Note: λ is the Poisson rate.

Poisson Parameter Scaling

If events occur according to a Poisson process with rate λ, then the rate for a Poisson process in an interval of length t is λt. e.g. :

Every minute, cars pull up to a drive-through according to a Poisson process with rate λ = 3.

- In an interval of length 1 hour, the rate is $\lambda = 180$.

3 * 60 (minutes in an hour) $= 180$

Examples

2.6

2

Suppose Albert receives spam mail every day according to a Poisson Process with an expected value of 10.

What is the distribution of X?

What is the probability that Albert receives 8 items of spam in a given day?

What is the probability that Albert receives 10 items of spam in a given day?

2 Suppose Albert receives spam mail every day according to a Poisson Process with an expected value of 10.

Find P[Albert receives 10 items of spam in a given day].

Find P[0 items of spam in a given day]?

Find P[Albert receives 1 item of spam in a given hour]?