Discrete Distributions - Bernoulli and Binomial

David Armstrong

UCI

Distributions - General Form

- Statisticians have been studying distributions for a long time.
- We are trying to find the probability of seeing a specific observation from a sample, given some population parameters.
- We say $X \sim Distribution(\Phi)$
- We define Φ as representing the population parameters.

Our first look at a distribution.

- Set $X = \{0, 1\}$.
- Let $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- Calculate the expectation of X.

- Calculate the variance of X.

The Bernoulli Distribution

- Assume we perform one event.
- If we set $X = \{0, 1\}$.
- Let p be the probability of success.
- Let $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- Then we say X follows a Bernoulli distribution with parameter p
- Denoted $X \sim Bernoulli(p)$
- $f(x) = P(X = x) = p^x(1-p)^{1-x}$
  - $\mathrm{E}(X) = p$.
  - $\mathrm{Var}(X) = p(1-p)$.

Example: Assume we perform an experiment with one trial where probability of success is 42%. Write out a relative frequency table for the event.

Binomial Distribution Motivation

- **Example:** Assume we perform an experiment with 4 repeated trials where the probability of success for each trial is p.
- What is the support of X?
- How many total outcomes are there?
- Draw out a table of all possible outcomes for each element in the support.
- Below each column, write probability of each element in the **Support**

The Binomial Distribution

- The binomial random variables with n trials and p parameter can be characterized as the number of success' in n independent trials.
- We define the probability of success on any given trial to be p, and the probability of no success on a given trial to be $1 - p$.
- The probability of each independent event does not change. p is constant
- Then we say X follows a Binomial distribution with parameters n and p.
- Denoted $X \sim Binomial(n, p)$
  - E$(X) = np$.
  - Var$(X) = np(1 - p)$.
- $f(x) = P(X = x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, 2, ..., n.$

R Code

- The pmf of the binomial is dbinom(x, n, p) = P(X=x).
- The cdf of the binomial is pbinom(x, n, p) = P(X≤x).

Examples of Binomial random variables.

- The number of heads to show up when flipping a fair coin 1000 times. ($n = 1000$ and $p = 0.5$).
- A company has 123 employees. All employees are independent of one another, and the probability that a single employee has certification is p. Then the number of employees that have certificates (out of the 123) is a binomial random variable.
- You roll a fair die 21650 times. The number of times a 3 shows is a binomial random variable. Here, $p = \frac{1}{6}$ is the probability of seeing a 3 on a given roll. The probability of not seeing a 3 (1,2,4,5, or 6) is $1 - p = \dfrac{5}{6}$.

Binomial Distribution - Set Values Example

Let X be a random variable that follows a binomial distribution with $n = 22$ and $p = \dfrac{1}{4}$

- What is the distribution of X

- Calculate the expectation of X.

- Calculate the variance of X.

- Calculate the standard deviation of X.

- What is the probability of seeing 5 successes?

- What is the probability that we see less than 5 successes?

- What is the probability that we see at most 3 successes?

- What is the probability that we see at least 4 successes?

Binomial Distribution - Wireless Households

Example According to CTIA, 32% of all U.S. housholds are wireless-only households, meaning they have no landline. In a random sample of 20 households, what is the probability that:

- Exactly 5 are wireless-only?
- Fewer that 3 are wireless-only?
- How many households in your sample would you expect to be wireless-only?
- What is the standard deviation of homes in your sample that would be wireless-only?