Chapter 4

Distribution[1]

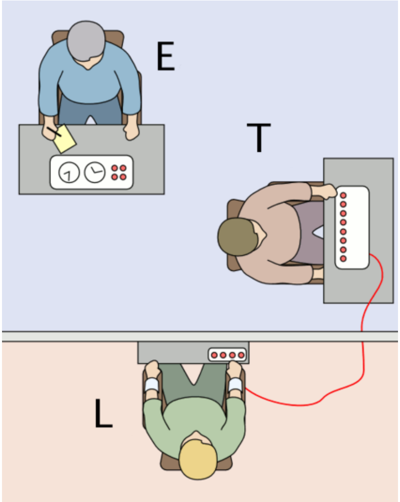
Department of Mathematics & Statistics
North Carolina A&T State University

[1] These notes use content from OpenIntro Statistics Slides by Mine Cetinkaya-Rundel.

Geometric distribution

Milgram experiment

- Stanley Milgram, a Yale University psychologist, conducted a series of experiments on obedience to authority starting in 1963.
- Experimenter(E) orders the teacher (T), the subject of the experiment, to give severe electric shocks to a learner (L) each time the learner answers a question incorrectly.
- The learner is actually an actor, and the electric shocks are not real, but a prerecorded sound is played each time the teacher administers an electric shock.



Milgram experiment

- These experiments measured the willingness of study participants to obey an authority figure who instructed them to perform acts that conflicted with their personal conscience.
- Milgram found that about 65% of people would obey authority and give such shocks.
- Over the years, additional research suggested this number is approximately consistent across communities and time.

Bernouilli random variables

- Each person in Milgram's experiment can be thought of as a **trial**.
- A person is labeled a **success** if she refuses to administer a severe shock, and **failure** if she administers such shock.
- Since only 35% of people refused to administer a shock, **probability of success** is $\mathbf{p = 0.35}$.
- When an individual trial has only two possible outcomes, it is called a **Bernoulli random variable**.

Geometric distribution

Dr. Smith wants to repeat Milgram's experiments but she only wants to sample people until she finds someone who will not inflict a severe shock. What is the probability that she stops after the first person?

$$\text{P}(1^{st} \text{ person refuses}) = 0.35$$

Geometric distribution

Dr. Smith wants to repeat Milgram's experiments but she only wants to sample people until she finds someone who will not inflict a severe shock. What is the probability that she stops after the first person?

$$\text{P}(1^{st} \text{ person refuses}) = 0.35$$

... the third person?

$$\text{P}(1^{st} \text{ and } 2^{nd} \text{ shock, } 3^{rd} \text{ refuses})$$
$$= \frac{S}{0.65} \times \frac{S}{0.65} \times \frac{R}{0.35} = 0.65^2 \times 0.35 \approx 0.15$$

Geometric distribution

Dr. Smith wants to repeat Milgram's experiments but she only wants to sample people until she finds someone who will not inflict a severe shock. What is the probability that she stops after the first person?

$$\text{P}(1^{st} \text{ person refuses}) = 0.35$$

... the third person?

$$\text{P}(1^{st} \text{ and } 2^{nd} \text{ shock, } 3^{rd} \text{ refuses})$$
$$= \frac{S}{0.65} \times \frac{S}{0.65} \times \frac{R}{0.35} = 0.65^2 \times 0.35 \approx 0.15$$

... the tenth person?

Geometric distribution

Dr. Smith wants to repeat Milgram's experiments but she only wants to sample people until she finds someone who will not inflict a severe shock. What is the probability that she stops after the first person?

$$\text{P}(1^{st} \text{ person refuses}) = 0.35$$

... the third person?

$$\text{P}(1^{st} \text{ and } 2^{nd} \text{ shock, } 3^{rd} \text{ refuses})$$
$$= \frac{S}{0.65} \times \frac{S}{0.65} \times \frac{R}{0.35} = 0.65^2 \times 0.35 \approx 0.15$$

... the tenth person?

$$\text{P}(1 \text{ shocks, } 10^{th} \text{ refuses})$$
$$= \underbrace{\frac{S}{0.65} \dots \frac{S}{0.65}} \times \frac{R}{0.35} = 0.65^9 \times 0.35 \approx 0.0072$$

Geometric distribution

Geometric distribution describes the waiting time until a success for **independent and identically distributed (iid)** Bernoulli random variables.

- Independence: Outcomes of trials don't affect each other.
- Identical: The probability of success is the same for each trial.

Geometric distribution

Geometric distribution describes the waiting time until a success for **independent and identically distributed (iid)** Bernoulli random variables.

- Independence: Outcomes of trials don't affect each other.
- Identical: The probability of success is the same for each trial.

Geometric probabilities

If p represents probability of success, $(1-p)$ represents probability of failure, and n represents number of independent trials

$$\text{P(success on the } n^{th} \text{ trial)} = (1-p)^{n-1}p$$

Practice

Can we calculate the probability of rolling a 6 for the first time on the 6^{th} roll of a die using the geometric distribution? Note that what was a success (rolling a 6) and what was a failure (not rolling a 6) are clearly defined and one or the other must happen for each trial.

A) No, on the the roll of a die there are more than 2 possible outcomes.

B) Yes, why not?

Practice

Can we calculate the probability of rolling a 6 for the first time on the 6^{th} roll of a die using the geometric distribution? Note that what was a success (rolling a 6) and what was a failure (not rolling a 6) are clearly defined and one or the other must happen for each trial.

A) No, on the the roll of a die there are more than 2 possible outcomes.

B) **Yes, why not?**

$$\text{P(6 on the } 6^{th} \text{ roll)} = (\tfrac{5}{6})^5 \cdot (\tfrac{1}{6}) \approx 0.067$$

Expected value

How many people is Dr. Smith expected to test before finding the first one that refuses to administer the shock?

Expected value

How many people is Dr. Smith expected to test before finding the first one that refuses to administer the shock?

The expected value, or the mean, of a geometric distribution is defined as $\frac{1}{p}$.

$$\mu = \frac{1}{p} = \frac{1}{0.35} = 2.86$$

Expected value

How many people is Dr. Smith expected to test before finding the first one that refuses to administer the shock?

The expected value, or the mean, of a geometric distribution is defined as $\frac{1}{p}$.

$$\mu = \frac{1}{p} = \frac{1}{0.35} = 2.86$$

She is expected to test 2.86 people before finding the first one that refuses to administer the shock.

Expected value

How many people is Dr. Smith expected to test before finding the first one that refuses to administer the shock?

The expected value, or the mean, of a geometric distribution is defined as $\frac{1}{p}$.

$$\mu = \frac{1}{p} = \frac{1}{0.35} = 2.86$$

She is expected to test 2.86 people before finding the first one that refuses to administer the shock.

But how can she test a non-whole number of people?

Expected value and its variability

Mean and standard deviation of geometric distribution

$$\mu = \frac{1}{p} \qquad\qquad \sigma = \sqrt{\frac{1-p}{p^2}}$$

Expected value and its variability

Mean and standard deviation of geometric distribution

$$\mu = \frac{1}{p} \qquad \sigma = \sqrt{\frac{1-p}{p^2}}$$

- Going back to Dr. Smith's experiment:

$$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-0.35}{0.35^2}} = 2.3$$

Expected value and its variability

Mean and standard deviation of geometric distribution

$$\mu = \frac{1}{p} \qquad\qquad \sigma = \sqrt{\frac{1-p}{p^2}}$$

- Going back to Dr. Smith's experiment:

$$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-0.35}{0.35^2}} = 2.3$$

- Dr. Smith is expected to test 2.86 people before finding the first one that refuses to administer the shock, give or take 2.3 people.

Expected value and its variability

Mean and standard deviation of geometric distribution

$$\mu = \frac{1}{p} \qquad\qquad \sigma = \sqrt{\frac{1-p}{p^2}}$$

- Going back to Dr. Smith's experiment:

$$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-0.35}{0.35^2}} = 2.3$$

- Dr. Smith is expected to test 2.86 people before finding the first one that refuses to administer the shock, give or take 2.3 people.
- These values only make sense in the context of repeating the experiment many many times.