M6S1 - Statistical Hypotheses

Professor Jarad Niemi

STAT 226 - Iowa State University

October 23, 2018

Outline

- Statistical Modeling:
    - Independent
    - Identically distributed
    - Normal
    - Parameters
- Statistical Hypotheses
    - Scientific hypotheses
    - Statistical hypotheses
    - Null vs alternative hypotheses
    - One-sided vs two-sided

Confidence interval construction

The United State Department of Agriculture National Agricultural Statistics Service reports the estimated corn yield in Iowa every year. To do so, they survey a random sample of corn growers and ask those growers to report the mean yield per acre on their farm. In 2017, the 110 surveyed growers had an average yield of 202.0 bushels per acre with a standard deviation of 31.6 bushels per acre. Construct a 95% confidence interval for the mean corn yield across Iowa.

Let X_i be the mean yield on farm i with $E[X_i] = \mu$ and $SD[X_i] = \sigma$ which are both unknown. We had a sample size of 110 with $\overline{x} = 202.2$ bushels per acre and $s = 31.6$ bushels per acre. With a confidence level of 95%, we have a significance level of 0.05, and a critical value of $t_{109,0.025} < t_{100,0.025} = 1.984$. Thus a 95% confidence interval for the mean yield across growers is

$$202 \pm 1.984 \frac{31.6}{\sqrt{110}} = (196.3 \text{ bushels per acre}, 207.7 \text{ bushels per acre}).$$

Assumptions

Let X_i be the mean yield on farm i and assume

$$X_i \stackrel{iid}{\sim} N(\mu, \sigma^2).$$

where iid stands for independent and identically distributed. We are assuming

- X_i are independent,
- X_i are identically distributed, i.e. each X_i is $N(\mu, \sigma^2)$,
- X_i are normally distributed, and
- X_i have a common mean μ and standard deviation σ.

Independence

Recall that X_1 is statistically independent of X_2 if the value of X_1 does not affect the distribution of X_2. In the corn yield example, $X_2 \sim N(\mu, \sigma^2)$, but suppose I told you that one farm had a yield of 210 bushels per acre. Does that change the distribution of X_2?

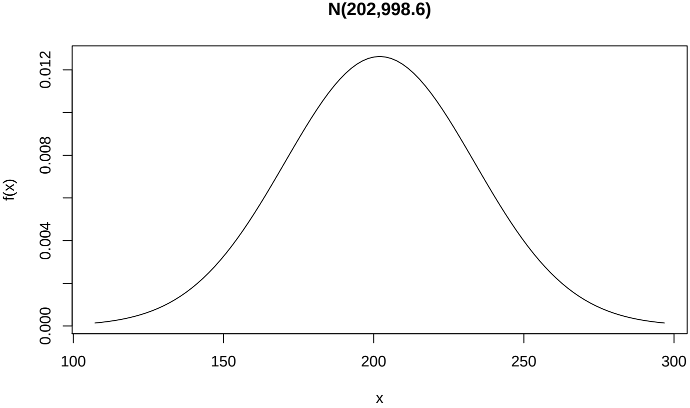
Common ways for independence to be violated:

- Temporal effects, e.g. yield this year is likely similar to yield last year
- Spatial effects, e.g. yield nearby is probably similar
- Clustering, e.g. these growers all used the same corn variety

Everything we do in this class requires the independence assumption, but you should be aware that it may violated easily.

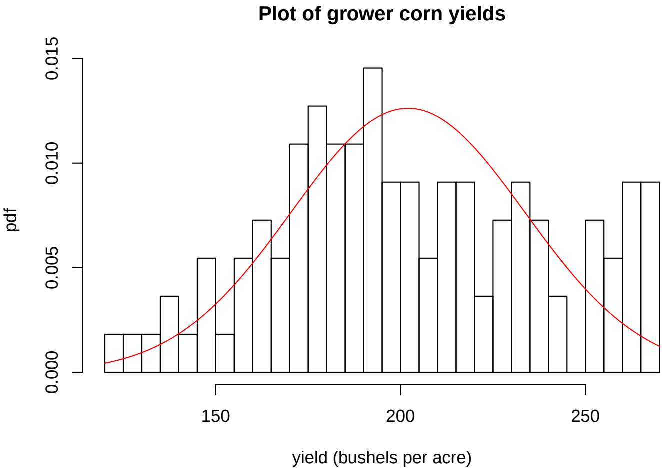
Identically distributed

Identically distributed means that each random variable has the same distribution, e.g. $X_i \sim N(\mu, \sigma^2)$ means that each X_i has a normal distribution with mean μ and standard deviation σ.



Normal

We can plot a histogram of the data to determine whether it is approximately normal.



Robustness

Typically none of our assumptions are met exactly. But the t-tools, e.g. confidence intervals based on the t distribution, are pretty robust to deviations from these assumptions. I would focus on lack of independence, e.g.

- temporal effects,
- spatial effects, and
- clustering.

A random sample will go a long way to help ensure that your data are independent.

Parameters

Recall that

- μ is the population mean and
- σ is the population standard deviation.

We've assumed each observation has the same mean and standard deviation. Often we would like to make formal statements about these parameters (typically the mean), e.g.

- The mean corn yield in Iowa is greater than 200 bushels per acre.
- The mean corn yield in Iowa is greater than last year.
- The mean corn yield in Iowa is different than last year.
- The mean corn yield in Iowa is less than last year.

To make these formal statements about a population parameter, we turn to Statistical Hypotheses.

Scientific Hypotheses

A scientific hypothesis is a statement about how we think the world may work.

Here are some scientific hypotheses that we may be interested in testing

- The coin is biased.
- Subway's chicken breast is less than half chicken.
- Average human body temperature is 98.6^oF.
- Corn yield is higher when fertilizer is added.
- High doses of vitamin C help prevent illness (or reduce illness duration).
- Training at least 10 hours a week helps prevent injury.
- An advertising strategy increased sales.

Statistical hypotheses

Statistical hypotheses are statements about the model assumptions. In this course, they will always be statements about the population parameters, specifically the population mean. Examples:

- Let X_i be an indicator the ith coin flipped heads with $E[X_i] = p$. An unbiased coin has $p = 0.5$ and a biased coin has $p \neq 0.5$.
- Let X_i be the percentage of chicken in breast i with $E[X_i] = \mu$. If $\mu < 50\%$, then (on average) the chicken breasts are less than half chicken.
- Let X_i be the body temperature for individual i with $E[X_i] = \mu$. If $\mu = 98.6^o$F, then the average human body temperature is 98.6^oF and $\mu \neq 98.6^o$F otherwise.

The hypotheses are always about the population and never about an individual.

Null vs alternative hypotheses

The methodology we will use (based on p-values) requires us to specify a null hypothesis and an alternative hypothesis.

Definition

The null hypotheses, H_0, is the generally accepted (or default) state of the world. The alternative hypothesis, H_a, is a proposed deviation from the generally accepted (or default) state of the world.

Examples:

- Coin flipping: $H_0 : p = 0.5$ versus $H_a : p \neq 0.5$.
- Subway: $H_0 : \mu \geq 50\%$ versus $H_a : \mu < 50\%$.
- Temperature: $H_0 : \mu = 98.6^o$F versus $H_a : \mu \neq 98.6^o$F.

The null hypothesis always includes the equality and, typically, we ignore the inequality, e.g.

- Subway: $H_0 : \mu = 50\%$ versus $H_a : \mu < 50\%$.

One-sided vs two-sided hypotheses

Definition

A one-sided alternative hypothesis has an inequality, i.e. $<$ or $>$, is is associated with the scientific hypotheses that include the words *less than* or *greater than*. A two-sided alternative hypothesis has a *not equal to sign*, i.e. $\neq$ and is associated with the scientific hypotheses that does not specify a direction.

Examples:

- Coin flipping: two-sided $H_0 : p = 0.5$ versus $H_a : p \neq 0.5$.
- Subway: one-sided $H_0 : \mu \geq 50\%$ versus $H_a : \mu < 50\%$.
- Temperature: two-sided $H_0 : \mu = 98.6^o$F versus $H_a : \mu \neq 98.6^o$F.

ACT scores

The mean composite score on the ACT among the students at a large Midwestern University is 24. We wish to know whether the average composite ACT score for business majors is different from the average for the University. We sample 100 business majors and calculate an average score of 26 with a standard deviation of 4.

Let X_i be the composite ACT score for business student i with $E[X_i] = \mu$. We have a null hypothesis that the average composite ACT score for business students is 24 and two-sided alternative hypothesis. So we have

$$H_0 : \mu = 24 \quad \text{versus} \quad H_a : \mu \neq 24.$$

https://wiki.uiowa.edu/display/bstat/Hypothesis+Testing

Foothill Hosiery socks

Foothill Hosiery recently received an order for childrens socks decorated with embroidered patches of cartoon characters. Foothill did not have the right machinery to sew on the embroidered patches and contracted out the sewing. While the order was filled and Foothill made a profit on it, the sewing contractors price seemed high, and Foothill had to keep pressure on the contractor to deliver the socks by the date agreed upon. Foothills CEO, John McGrath, has explored buying the machinery necessary to allow Foothill to sew patches on socks themselves. He has discovered that if more than a quarter of the childrens socks they make are ordered with patches, the machinery will be a sound investment. John asks Kevin to find out if more than 35 percent of childrens socks are being sold with patches.

Let X_i be an indicator that sock i has patches with $E[X_i] = \mu$ (or p). We have an alternative hypothesis that more than 35 percent of socks have patches and a null hypothesis that is the opposite. So we have

$$H_0 : \mu \leq 0.35 \quad \text{versus} \quad H_a : \mu > 0.35$$

or

$$H_0 : \mu = 0.35 \quad \text{versus} \quad H_a : \mu > 0.35$$

https://opentextbc.ca/introductorybusinessstatistics/chapter/hypothesis-testing-2/