Inference part 1

STAT-S520

Arturo Valdivia

02-28-23

- These slides complement material from ISI Chapter 9

So far, we have a random sample $X_1, X_2, \ldots, X_n \overset{i.i.d}{\sim} \mathbb{P}$ with $EX_i = \mu$ and $VarX_i = \sigma^2$ for $i = 1, \ldots, n$ and

$$Y = \sum_{i=1}^{n} X_i$$

and

$$\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- If we know $\mathbb{P}$, we know everything about Y and $\bar{X}_n$.

The inference problem (in terms of μ)

- Statistical inference describes any procedure for extracting information about a probability distribution (population) from an observed sample. Let's focus, for example, in the population mean, μ
- We want to determine whether or not a claim about μ is correct or estimate as accurately as possible the value of μ

Types of statistical inference

- Point estimation
- Hypothesis testing
- Set estimation

Examples

- We believe drivers are speeding near a school zone. We take a random sample of the speed of cars passing by to determine what is their average speed limit and use it to conclude whether or not drivers are speeding on the school zone.

Examples

- We believe drivers are speeding near a school zone. We take a random sample of the speed of cars passing by to determine what is their average speed limit and use it to conclude whether or not drivers are speeding on the school zone.
- We do not believe on psychic powers but someone claims to be a psychic. To test this, we use a sign that randomly shows left or right, and ask the prospective psychic to guess whether the sign is showing left of right. We do this 20 times and use this information to determine whether the individual is a psychic.

Examples

- We believe drivers are speeding near a school zone. We take a random sample of the speed of cars passing by to determine what is their average speed limit and use it to conclude whether or not drivers are speeding on the school zone.
- We do not believe on psychic powers but someone claims to be a psychic. To test this, we use a sign that randomly shows left or right, and ask the prospective psychic to guess whether the sign is showing left of right. We do this 20 times and use this information to determine whether the individual is a psychic.
- I would like to know what time the entire population of STAT-S 520 students go to sleep, on average. I randomly select a sample of students and find the sample average of sleeping times.

Point Estimation: Terminology

- Estimand: The value we are trying to estimate
- Estimator: The method or process used to estimate a parameter
- Estimate: The observed value obtained using the method

Example 1

Point Estimation

Point estimation is a method used to obtained a single value as estimator or estimate

Properties of Estimators

- **Unbiasedness** An estimator is unbiased if its expected value is equal to the estimand

- **Consistency** An estimator is consistent if it converges in probability to the estimand

Example 2

- The plug-in estimate for the variance is

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)^2}{n}$$

The plug-in estimator is biased:

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] = \frac{n-1}{n}\sigma^2$$

- The sample variance is the estimate

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

with corresponding estimator

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

Example 2 Simulation

```
mu = 0; sigma2 = 25; sigma = sqrt(sigma2); n = 60
plug.sd = function(x) sqrt(1/n*sum((x - mean(x))^2))
set.seed(1003)
vec_plug_sd =replicate(10^4, plug.sd(rnorm(n, mu, sigma)))
mean(vec_plug_sd)
```

```
#> [1] 4.949723
```

```
set.seed(1003)
vec_sample_sd =replicate(10^4, sd(rnorm(n, mu, sigma)))
mean(vec_sample_sd)
```

```
#> [1] 4.991494
```

Hypothesis Testing

- A claim about μ is made (null hypothesis) about the mean of the population
- A random sample from this population is obtained and used to obtain the sample mean $\bar{x}_n$
- Assuming the claim about μ is correct, we determine how likely is to observe a value as extreme as or more extreme than $\bar{x}_n$
  - If, getting $\bar{x}_n$ is very unlikely, we reject the claim (reject the null hypothesis)

Hypotheses

- The null hypothesis, H_0, is what we initially assume to be true. It's a statement about μ.
  - For procedural reasons, we **always** include the equal sign in the statement under H_0
- The alternative hypothesis, H_1, is what we would conclude (about μ) if we were to reject H_0.

Examples

Test Statistic

The statistic is the method we use to assess the claim made under the null hypothesis. Many statistics exist, the one we encounter often looks like this:

$$statistic = \frac{estimator - parameter\ under\ H_0}{standard\ error}$$

Hypothesis about μ when σ is known

When the hypothesis is made about μ, and σ is known, the test statistic is given by

$$Z = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$

and due to the CLT, $Z \sim N(0, 1)$.

Hypothesis about μ when σ is unknown

When σ is unknown (as in most real-life problems), the test statistic is given by

$$T = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}$$

where $S_n = \sqrt{\frac{\sum(X_i - \bar{X}_n)^2}{n-1}}$ is the sample standard deviation, another estimator. Given the added uncertainty of S_n, T is no longer normal, but T follows a T-distribution with $n-1$ degrees of freedom and we write $T \sim T_{n-1}$

Observed test statistic

Once you collect a (random) sample of n observations, the sample observed is $\overrightarrow{x} = (x_1, \ldots, x_n)$. If σ is known, the observed test statistic is

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

and if σ is unknown, the observed test statistic is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

where $s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$ is the sample standard deviation.

Significance Probability (p-value)

- The p-value is the probability of observing a test statistic that is as extreme or more extreme than the one observed in our data, when we assume that H_0 is true.
- The p-value depends on the hypotheses statements made

Conclusion and Interpretation of results

- We want to reject H_0 if the observed estimate is highly unlikely to have been obtained by chance
  - The smaller the p-value, the more evidence to reject H_0.
- While how small p-value needs to be is somewhat arbitrary, it should be guided by how important is not to make a mistake by rejecting H_0 when it is actually true (Type I Error) or by failing to reject H_0 when it is actually false (Type II Error).

Reasonable ranges for the *p*-value

To guide you in the decision process, some reasonable values (although still arbitrary) can be:

- If p-value > 0.1, do not reject H_0.
- If p-value < 0.001, reject H_0.
- If $0.001 \leq p$-value ≤ 0.1, decide based on your own perception of this uncertainty (different people may make different decisions).
- Alternatively, come up with a significance level, α, such that if p-value $\leq \alpha$, we reject H_0, and if p-value $> \alpha$, we fail to reject H_0.
  - The choice of α should be set before collecting and/or observing the random sample.