YSU ASDS, Statistics, Fall 2019

Lecture 22

Michael Poghosyan

11 Nov 2019

Contents

Last Lecture ReCap

- Give the definition of the $(1-\alpha)$-level CI.

Last Lecture ReCap

- Give the definition of the $(1-\alpha)$-level CI.
- Give the Chebyshev Inequality.

Last Lecture ReCap

- Give the definition of the $(1 - \alpha)$-level CI.
- Give the Chebyshev Inequality.
- What is the **R** command to generate 20 random numbers from the *Cauchy*(2) distribution?

Last Lecture ReCap

- Give the definition of the $(1-\alpha)$-level CI.
- Give the Chebyshev Inequality.
- What is the **R** command to generate 20 random numbers from the *Cauchy*(2) distribution?
- Give a $(1-\alpha)$-level CI for *p* in *Bernoulli*(*p*) Model.

CI by Pivotal Quantity Method

Again, we want to construct a CI of CL $1-\alpha$ for θ, using the Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta.$$

CI by Pivotal Quantity Method

Again, we want to construct a CI of CL $1-\alpha$ for θ, using the Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta.$$

Suppose we can construct a function of $X_1, X_2, ..., X_n$ and our unknown parameter θ,

$$g(X_1, X_2, ..., X_n, \theta)$$

such that

CI by Pivotal Quantity Method

Again, we want to construct a CI of CL $1-\alpha$ for θ, using the Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta.$$

Suppose we can construct a function of $X_1, X_2, ..., X_n$ and our unknown parameter θ,

$$g(X_1, X_2, ..., X_n, \theta)$$

such that

- the distribution of $g(X_1, ..., X_n, \theta)$ is independent of θ;

CI by Pivotal Quantity Method

Again, we want to construct a CI of CL $1 - \alpha$ for θ, using the Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta.$$

Suppose we can construct a function of $X_1, X_2, ..., X_n$ and our unknown parameter θ,

$$g(X_1, X_2, ..., X_n, \theta)$$

such that

- the distribution of $g(X_1, ..., X_n, \theta)$ is independent of θ;
- g is invertible in terms of θ, say, is strictly monotonic in θ

CI by Pivotal Quantity Method

Again, we want to construct a CI of CL $1-\alpha$ for θ, using the Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta.$$

Suppose we can construct a function of $X_1, X_2, ..., X_n$ and our unknown parameter θ,

$$g(X_1, X_2, ..., X_n, \theta)$$

such that

- the distribution of $g(X_1, ..., X_n, \theta)$ is independent of θ;
- g is invertible in terms of θ, say, is strictly monotonic in θ

Then we call $g(X_1, ..., X_n, \theta)$ to be a **Pivot** for our model.

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

- Find a and b such that

$$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) \geq 1 - \alpha;$$

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

- Find a and b such that

  $$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) \geq 1 - \alpha;$$

  Because of the first condition, a and b will not depend on θ;

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

- Find a and b such that
$$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) \geq 1 - \alpha;$$
Because of the first condition, a and b will not depend on θ;
- now solve the inequality
$$a < g(X_1, ..., X_n, \theta) < b$$
for θ for the obtained a and b.

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

- Find a and b such that

$$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) \geq 1 - \alpha;$$

  Because of the first condition, a and b will not depend on θ;

- now solve the inequality

$$a < g(X_1, ..., X_n, \theta) < b$$

  for θ for the obtained a and b.

This will give some inequality of the form $L < \theta < U$, and hence,

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha,$$

so we can take (L, U) as a CI.

CI by Pivotal Quantity Method, Cont'd

How to use Pivots to construct CIs? The general idea is:

- Find a and b such that

  $$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) \geq 1 - \alpha;$$

  Because of the first condition, a and b will not depend on θ;

- now solve the inequality

  $$a < g(X_1, ..., X_n, \theta) < b$$

  for θ for the obtained a and b.

This will give some inequality of the form $L < \theta < U$, and hence,

$$\mathbb{P}(L < \theta < U) \geq 1 - \alpha,$$

so we can take (L, U) as a CI.

Note: Usually, we are solving

$$\mathbb{P}\Big(a < g(X_1, X_2, ..., X_n, \theta) < b\Big) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ. So we will use it to obtain a Pivot.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ. So we will use it to obtain a Pivot.

We know that, since $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, then

$$\overline{X} \sim$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ. So we will use it to obtain a Pivot.

We know that, since $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, then

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ. So we will use it to obtain a Pivot.

We know that, since $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, then

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

This shows that $\overline{X}$ cannot serve as a Pivot, since its Distribution depends on the unknown μ.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is known. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: We know that $\overline{X}$ is a good Estimator for μ. So we will use it to obtain a Pivot.

We know that, since $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, then

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

This shows that $\overline{X}$ cannot serve as a Pivot, since its Distribution depends on the unknown μ.

But let us consider $\overline{X} - \mu$.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ!

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ! Uraaa! We found a Pivoooot.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ! Uraaa! We found a Pivoooot. In fact, usually people take the following Pivot:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ! Uraaa! We found a Pivoooot. In fact, usually people take the following Pivot:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

since working with $\mathcal{N}(0, 1)$ is easier: we have Standard Tables for Standard Normal!

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ! Uraaa! We found a Pivoooot. In fact, usually people take the following Pivot:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

since working with $\mathcal{N}(0, 1)$ is easier: we have Standard Tables for Standard Normal! So what??? Who cares for that in this digital era of the 4th Industrial Revolution? $\ddot{\smile}$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Clearly,

$$\overline{X} - \mu \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

So the Distribution of $\overline{X} - \mu$ is *independent of* μ! Uraaa! We found a Pivoooot. In fact, usually people take the following Pivot:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

since working with $\mathcal{N}(0, 1)$ is easier: we have Standard Tables for Standard Normal! So what??? Who cares for that in this digital era of the 4th Industrial Revolution? $\ddot\smile$

The name of the above ratio is Z-**statistics**, and we will meet this again in Hypotheses testing part.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals. But let us choose the *minimum length* interval, which will be

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals. But let us choose the *minimum length* interval, which will be a symmetric interval around 0 (prove!).

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0,1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals. But let us choose the *minimum length* interval, which will be a symmetric interval around 0 (prove!). So we need to find an inteval of the form $(-b, b)$ such that

$$\mathbb{P}(-b < Z < b) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0, 1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals. But let us choose the *minimum length* interval, which will be a symmetric interval around 0 (prove!). So we need to find an inteval of the form $(-b, b)$ such that

$$\mathbb{P}(-b < Z < b) = 1 - \alpha.$$

Because of the symmetry, we need to have that the area of both tails is $\alpha/2$:

$$\mathbb{P}(Z \leq -b) = \mathbb{P}(Z \geq b) = \frac{\alpha}{2},$$

hence,

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

Now, let us use this Pivot: we have $Z \sim \mathcal{N}(0,1)$. Let us find numbers a and b such that

$$\mathbb{P}(a < Z < b) = 1 - \alpha.$$

This means that we want to select an interval, over which the area under the Standard Normal PDF is $1 - \alpha$. Clearly, there exist infinitely many this kind of intervals. But let us choose the *minimum length* interval, which will be a symmetric interval around 0 (prove!). So we need to find an inteval of the form $(-b, b)$ such that

$$\mathbb{P}(-b < Z < b) = 1 - \alpha.$$

Because of the symmetry, we need to have that the area of both tails is $\alpha/2$:

$$\mathbb{P}(Z \leq -b) = \mathbb{P}(Z \geq b) = \frac{\alpha}{2},$$

hence,

$$b = z_{1-\frac{\alpha}{2}},$$

where $z_{1-\alpha/2}$ is the $1 - \frac{\alpha}{2}$ quantile of the Standard Normal Distribution.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

So we obtained

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

So we obtained

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

We plug here the value of Z:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha,$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

So we obtained

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

We plug here the value of Z:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha,$$

and solve for μ:

$$\mathbb{P}\left(\overline{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is known, Pivotal Method

So we obtained

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

We plug here the value of Z:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha,$$

and solve for μ:

$$\mathbb{P}\left(\overline{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Hence, the following interval:

$$\left(\overline{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \ \overline{X} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

is a $(1 - \alpha)$-level CI for μ.

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known.

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known. We construct it using a Random Sample

$$X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known. We construct it using a Random Sample

$$X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Since $1 - \alpha = 0.95$, then $\alpha = 0.05$.

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known. We construct it using a Random Sample

$$X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Since $1 - \alpha = 0.95$, then $\alpha = 0.05$. By the above formula, we need to calculate the Standard Normal quantile $z_{1-\alpha/2} =$

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known. We construct it using a Random Sample

$$X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Since $1 - \alpha = 0.95$, then $\alpha = 0.05$. By the above formula, we need to calculate the Standard Normal quantile $z_{1-\alpha/2} = z_{0.975}$.

Example

Example: Assume we want to construct a 95% CI for μ in the $\mathcal{N}(\mu, \sigma^2)$ Model, when σ is given, known. We construct it using a Random Sample

$$X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Since $1 - \alpha = 0.95$, then $\alpha = 0.05$. By the above formula, we need to calculate the Standard Normal quantile $z_{1-\alpha/2} = z_{0.975}$.

R gives:

```r
qnorm(0.975)
```

```
## [1] 1.959964
```

so our 95% CI will be

$$\left(\overline{X} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, \;\; \overline{X} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right).$$

Example

Example with R: We generate random numbers from $\mathcal{N}(2.31, 4)$ (so here we assume we know the true parameter value of μ).

```
sigma <- 2
n <- 20
smpl <- rnorm(n, mean = 2.31, sd = sigma)
smpl
```

```
##  [1] -0.61517958  0.26141094 -0.47118696  3.42826333  1.
##  [6]  0.06724348  1.18915020  0.70609049  4.60922396  0.
## [11]  3.92192237 -1.71053629  3.07221264  2.21058725  5.
## [16]  4.43852249 -0.54471636  2.27825637  2.97237654 -1.
```

Example

Example with R: We generate random numbers from $\mathcal{N}(2.31, 4)$ (so here we assume we know the true parameter value of μ).

```
sigma <- 2
n <- 20
smpl <- rnorm(n, mean = 2.31, sd = sigma)
smpl
```

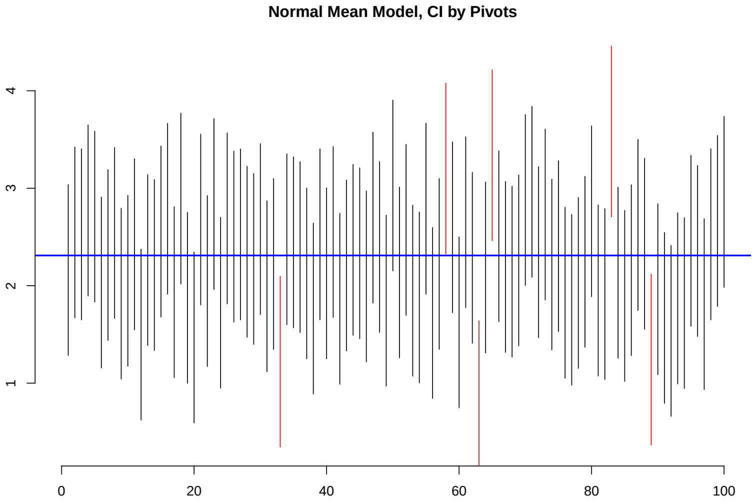
```
##  [1] -0.61517958  0.26141094 -0.47118696  3.42826333  1.
##  [6]  0.06724348  1.18915020  0.70609049  4.60922396  0.
## [11]  3.92192237 -1.71053629  3.07221264  2.21058725  5.
## [16]  4.43852249 -0.54471636  2.27825637  2.97237654 -1.
```

Now we construct the 95% CI by the above formula:

```
me <- 1.96* sigma/sqrt(n) #the Margin of Error
c(mean(smpl) - me, mean(smpl) + me)
```

```
## [1] 0.680515 2.433592
```

Example, Simulation



Example, Simulation, Code

```
mu <- 2.31; sigma <- 2
conf.level <- 0.95; a = 1 - conf.level
sample.size <- 20; no.of.intervals <- 100
z <- qnorm(1-a/2) ## our quantile
ME <- z*sigma/sqrt(sample.size) #Margin of Error

plot.new()
plot.window(xlim = c(0,no.of.intervals), ylim = c(mu-2,mu+2))
axis(1); axis(2)
title("Normal Mean Model, CI by Pivots")
for(i in 1:no.of.intervals){
  x <- rnorm(sample.size, mean = mu, sd = sigma)
  lo <- mean(x) - ME; up <- mean(x) + ME
  if(lo > mu || up < mu){
    segments(c(i), c(lo), c(i), c(up), col = "red")
  }
  else{
    segments(c(i), c(lo), c(i), c(up))
  }
}
abline(h = mu, lwd = 2, col = "blue")
```

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

In this case, again

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

In this case, again

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1),$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

In this case, again

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1),$$

but, unfortunately, we cannot use Z, since the result will contain σ, **which is unknown to us**.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2).$$

Assume σ^2 is **unknown**, which is more realistic. Given $\alpha \in (0, 1)$, we want to construct a CI of CL $1 - \alpha$ for μ, using a Pivot.

Solution: Again we start from $\overline{X}$, having that it is a good Estimator for μ. Again, from the fact that $X_k \sim \mathcal{N}(\mu, \sigma^2)$ are IID, we will have

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

In this case, again

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1),$$

but, unfortunately, we cannot use Z, since the result will contain σ, **which is unknown to us**. So we need to adjust Z.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

We know some good Estimators for σ: let us take, in this case, the following version of Sample SD:

$$S^2 = \frac{\sum_{k=1}^n (X_k - \overline{X})^2}{n-1}, \qquad i.e., \qquad S = \sqrt{\frac{\sum_{k=1}^n (X_k - \overline{X})^2}{n-1}}.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

We know some good Estimators for σ: let us take, in this case, the following version of Sample SD:

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}, \qquad i.e., \qquad S = \sqrt{\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}}.$$

Now, let us replace above σ with S:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

We know some good Estimators for σ: let us take, in this case, the following version of Sample SD:

$$S^2 = \frac{\sum_{k=1}^n (X_k - \overline{X})^2}{n-1}, \qquad i.e., \qquad S = \sqrt{\frac{\sum_{k=1}^n (X_k - \overline{X})^2}{n-1}}.$$

Now, let us replace above σ with S:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

OK, what's next? Is it a Pivot? Is the Distribution of t independent of μ?

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

We know some good Estimators for σ: let us take, in this case, the following version of Sample SD:

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}, \qquad i.e., \qquad S = \sqrt{\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}}.$$

Now, let us replace above σ with S:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

OK, what's next? Is it a Pivot? Is the Distribution of t independent of μ? Or, in general, do we know the Distribution of t ?

I know, not sure about you.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

We know some good Estimators for σ: let us take, in this case, the following version of Sample SD:

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}, \qquad i.e., \qquad S = \sqrt{\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}}.$$

Now, let us replace above σ with S:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

OK, what's next? Is it a Pivot? Is the Distribution of t independent of μ? Or, in general, do we know the Distribution of t ?

I know, not sure about you. Yet!

t-Distribution

It turns out that the Distribution of above

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

is the famous t-Distribution with $n - 1$ degrees of freedom:

[1] See, e.g., https://en.wikipedia.org/wiki/Student's_t-distribution

t-Distribution

It turns out that the Distribution of above

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

is the famous t-Distribution with $n - 1$ degrees of freedom:

Definition: If $X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2)$ are IID and

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n - 1},$$

then the Distribution of

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

is called the **Student's t-Distribution with $n - 1$ degrees of freedom**[1], and is denoted by $t(n - 1)$.

[1] See, e.g., https://en.wikipedia.org/wiki/Student's_t-distribution

Have you ever heard the term *Vechniy Student* ?

Have you ever heard the term *Vechniy Student* ? Here is he:

Have you ever heard the term *Vechniy Student* ? Here is he:

Have you ever heard the term *Vechniy Student* ? Here is he:

British statistician William Sealy Gosset, known as "Student"

Have you ever heard the term *Vechniy Student* ? Here is he:

British statistician William Sealy Gosset, known as "Student"

You can read about him at
https://en.wikipedia.org/wiki/William_Sealy_Gosset.

Student's Paper

Student's Paper

VOLUME VI MARCH, 1908 No. 1

BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

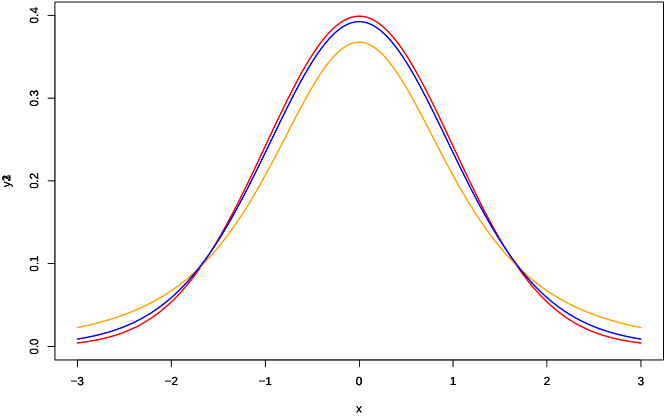
BY STUDENT.

Introduction.

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

t-Distribution

```
x <- seq(-3,3, 0.01)
y1 <- dnorm(x); y2 <- dt(x, df = 3); y3 <- dt(x, df = 15)
plot(x,y1, type = "l", col = "red", lwd = 2, ylim = c(0, 0.4))
par(new = T)
plot(x,y2, type = "l", col = "orange", lwd = 2, ylim = c(0, 0.4))
par(new = T)
plot(x,y3, type = "l", col = "blue", lwd = 2, ylim = c(0, 0.4))
```



CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ. We call it a t-**statistics**.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ. We call it a t-**statistics**.

Now, for $t \sim t(n-1)$, let us find numbers a and b such that

$$\mathbb{P}(a < t < b) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ. We call it a t-**statistics**.

Now, for $t \sim t(n-1)$, let us find numbers a and b such that

$$\mathbb{P}(a < t < b) = 1 - \alpha.$$

Again, because of the symmetry of the PDF of $t(n-1)$, the *minimum length* interval will be symmetric around 0.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ. We call it a t-**statistics**.

Now, for $t \sim t(n-1)$, let us find numbers a and b such that

$$\mathbb{P}(a < t < b) = 1 - \alpha.$$

Again, because of the symmetry of the PDF of $t(n-1)$, the *minimum length* interval will be symmetric around 0. So we need to find an inteval of the form $(-b, b)$ such that

$$\mathbb{P}(-b < t < b) = 1 - \alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Back to our Problem, we know that

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

and the Distribution of t is independent of μ, so it is a Pivot for μ. We call it a t-**statistics**.

Now, for $t \sim t(n-1)$, let us find numbers a and b such that

$$\mathbb{P}(a < t < b) = 1 - \alpha.$$

Again, because of the symmetry of the PDF of $t(n-1)$, the *minimum length* interval will be symmetric around 0. So we need to find an inteval of the form $(-b, b)$ such that

$$\mathbb{P}(-b < t < b) = 1 - \alpha.$$

Again, because of the symmetry, we need to have that the area of both tails is $\alpha/2$:

$$\mathbb{P}(t \leq -b) = \mathbb{P}(t \geq b) = \frac{\alpha}{2}.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

$$b = t_{1-\frac{\alpha}{2}}(n-1) = t_{n-1,1-\frac{\alpha}{2}},$$

where $t_{n-1,1-\alpha/2}$ is the $1 - \frac{\alpha}{2}$ quantile of the $t(n-1)$ Distribution.

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

$$b = t_{1-\frac{\alpha}{2}}(n-1) = t_{n-1,1-\frac{\alpha}{2}},$$

where $t_{n-1,1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the $t(n-1)$ Distribution.

So we obtained

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}\right) = 1-\alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

$$b = t_{1-\frac{\alpha}{2}}(n-1) = t_{n-1,1-\frac{\alpha}{2}},$$

where $t_{n-1,1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the $t(n-1)$ Distribution.

So we obtained

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}\right) = 1-\alpha.$$

We plug here the value of t:

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{n-1,1-\alpha/2}\right) = 1-\alpha,$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

$$b = t_{1-\frac{\alpha}{2}}(n-1) = t_{n-1,1-\frac{\alpha}{2}},$$

where $t_{n-1,1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the $t(n-1)$ Distribution.

So we obtained

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}\right) = 1-\alpha.$$

We plug here the value of t:

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{n-1,1-\alpha/2}\right) = 1-\alpha,$$

and solve for μ:

$$\mathbb{P}\left(\overline{X} - t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}}\right) = 1-\alpha.$$

CI for the Mean of $\mathcal{N}(\mu, \sigma^2)$, σ is **unknown**, PivMe

Hence,

$$b = t_{1-\frac{\alpha}{2}}(n-1) = t_{n-1,1-\frac{\alpha}{2}},$$

where $t_{n-1,1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the $t(n-1)$ Distribution.

So we obtained

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}\right) = 1-\alpha.$$

We plug here the value of t:

$$\mathbb{P}\left(-t_{n-1,1-\alpha/2} < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{n-1,1-\alpha/2}\right) = 1-\alpha,$$

and solve for μ:

$$\mathbb{P}\left(\overline{X} - t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}}\right) = 1-\alpha.$$

Hence, the following interval:

$$\left(\overline{X} - t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}},\ \ \overline{X} + t_{n-1,1-\alpha/2}\cdot\frac{S}{\sqrt{n}}\right)$$

is a $(1-\alpha)$-level CI for μ.

CI for μ, Normal Model, Notes

Note: To compare:

- If σ is known, $(1-\alpha)$-level CI for μ is
$$\overline{X} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$
- If σ is unknown, $(1-\alpha)$-level CI for μ is
$$\overline{X} \pm t_{n-1,1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

CI for μ, Normal Model, Notes

Note: To compare:

- If σ is known, $(1-\alpha)$-level CI for μ is

$$\overline{X} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

- If σ is unknown, $(1-\alpha)$-level CI for μ is

$$\overline{X} \pm t_{n-1,1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

Note: If we will compare the quantiles of the same level of $\mathcal{N}(0,1)$ with $t(n-1)$, we will see that CIs for the case when σ is unknown are wider than for the case when σ is known. This is intuitive, of course - to compensate the uncertainty in σ, we need to take a wider interval:

```
c(qnorm(0.975), qt(0.975, df = 3), qt(0.975, df = 20))
```

```
## [1] 1.959964 3.182446 2.085963
```

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows the time spent solving a hw for a (randomly chosen) student.

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows the time spent solving a hw for a (randomly chosen) student. Our unknown, μ is

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows the time spent solving a hw for a (randomly chosen) student. Our unknown, μ is the average time to solve a hw.

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows the time spent solving a hw for a (randomly chosen) student. Our unknown, μ is the average time to solve a hw. And we have an observation from a Random Sample

$$X_1, X_2, ..., X_{10} \sim \mathcal{N}(\mu, \sigma^2).$$

Example

Example: Assume we want to estimate the average time μ our Stat students are spending for a homework. We collect information from 10 students for the last hw, and get the following responses (in hrs):

```
[1] 2.17 1.42 2.13 0.56 1.21 2.22 1.35 2.37 1.47 1.70
```

Now we assume that the time spent for solving a hw is Normally Distributed with the Mean μ and some variance σ^2.

So $X \sim \mathcal{N}(\mu, \sigma^2)$ shows the time spent solving a hw for a (randomly chosen) student. Our unknown, μ is the average time to solve a hw. And we have an observation from a Random Sample

$$X_1, X_2, ..., X_{10} \sim \mathcal{N}(\mu, \sigma^2).$$

We construct a 95% CI for μ, the average time to solve the hw, by the above formula:

Example

```r
n <- 10
cl <- 0.95; a = 1-cl
t <- qt(1-a/2, df = n-1)
smpl<-c(2.17,1.42,2.13,0.56,1.21,2.22,1.35,2.37,1.47,1.70)
s <- sd(smpl)
me <-  t*s/sqrt(n) #the Margin of Error
c(mean(smpl) - me, mean(smpl) + me)
```

```
## [1] 1.253748 2.066252
```

Example

```
n <- 10
cl <- 0.95; a = 1-cl
t <- qt(1-a/2, df = n-1)
smpl<-c(2.17,1.42,2.13,0.56,1.21,2.22,1.35,2.37,1.47,1.70)
s <- sd(smpl)
me <-  t*s/sqrt(n) #the Margin of Error
c(mean(smpl) - me, mean(smpl) + me)
```

```
## [1] 1.253748 2.066252
```

Veeery unsatisfactory result, of course! Please spend **more time for your hw, read textbooks**!

Example, cont'd

Later, we will talk about the *t*-**Test**, let me now just do a *t*-Test for the hw solving hours:

```r
smpl <- c(2.17,1.42,2.13,0.56,1.21,2.22,1.35,2.37,1.47,1.70)
t.test(smpl)
```

```
##
##  One Sample t-test
##
## data:  smpl
## t = 9.2435, df = 9, p-value = 6.86e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.253748 2.066252
## sample estimates:
## mean of x
##      1.66
```

Example, cont'd

We can separate here the CI:

```
smpl <- c(2.17,1.42,2.13,0.56,1.21,2.22,1.35,2.37,1.47,1.70)
tst <- t.test(smpl) #Keeping the test result in tst
tst$conf.int
```

```
## [1] 1.253748 2.066252
## attr(,"conf.level")
## [1] 0.95
```

CI for μ, Normal Model, Summary

Let us summarize what we have obtained for this model.

CI for μ, Normal Model, Summary

Let us summarize what we have obtained for this model. The problem is: given a Random Sample

$$X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2),$$

and $\alpha \in (0, 1)$, we want to construct an $1 - \alpha$-level CI for the unknown parameter μ.

CI for μ, Normal Model, Summary

Let us summarize what we have obtained for this model. The problem is: given a Random Sample

$$X_1, X_2, ..., X_n \stackrel{IID}{\sim} \mathcal{N}(\mu, \sigma^2),$$

and $\alpha \in (0, 1)$, we want to construct an $1 - \alpha$-level CI for the unknown parameter μ.

We have considered 2 cases: when σ^2 was known and unknown. Here we give the summary:

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:		

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$
Distr. of Pivo:		

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$
Distr. of Pivo:	$Z \sim \mathcal{N}(0,1)$	$t \sim t(n-1)$
Reg w/prob $1-\alpha$		

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$
Distr. of Pivo:	$Z \sim \mathcal{N}(0,1)$	$t \sim t(n-1)$
Reg w/prob $1-\alpha$	$-z_{1-\alpha/2} < Z < z_{1-\alpha/2}$	$-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}$
CI for μ:		

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$
Distr. of Pivo:	$Z \sim \mathcal{N}(0,1)$	$t \sim t(n-1)$
Reg w/prob $1-\alpha$	$-z_{1-\alpha/2} < Z < z_{1-\alpha/2}$	$-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}$
CI for μ:	$\overline{X} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	$\overline{X} \pm t_{n-1,1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$

CI for μ, Normal Model, Summary

	σ^2 is known	σ^2 is unknown
Pivot:	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$
Distr. of Pivo:	$Z \sim \mathcal{N}(0,1)$	$t \sim t(n-1)$
Reg w/prob $1-\alpha$	$-z_{1-\alpha/2} < Z < z_{1-\alpha/2}$	$-t_{n-1,1-\alpha/2} < t < t_{n-1,1-\alpha/2}$
CI for μ:	$\overline{X} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	$\overline{X} \pm t_{n-1,1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$

Here S is the Sample Standard Deviation given by

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot?

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot? Yes, of course.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot? Yes, of course. What is the Distirbution of this r.v.?

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot? Yes, of course. What is the Distirbution of this r.v.? No, of course $\ddot\smile$.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot? Yes, of course. What is the Distirbution of this r.v.? No, of course ⌣̈. Well, we can use our Prob knowledge, but let's keep this to you.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **known**. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

We have $X_k \sim \mathcal{N}(\mu, \sigma^2)$, so

$$\frac{X_k - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

OK, the Distrib of the ratio is independent on the unknown Parameter, so this can serve as a Pivot. But please note that our Parameter was σ^2, for σ. We can take then the following:

$$\left(\frac{X_k - \mu}{\sigma}\right)^2.$$

Is this a Pivot? Yes, of course. What is the Distirbution of this r.v.? No, of course $\ddot\smile$. Well, we can use our Prob knowledge, but let's keep this to you. The fact is that this Pivot will not give a good result, as we are not using all the information we have.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

So we will use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 .$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

So we will use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 .$$

To show that this can serve as a Pivot, and to use it, we need to find the Distribution of Y.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

So we will use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2.$$

To show that this can serve as a Pivot, and to use it, we need to find the Distribution of Y. Fortunately, someone already found the Distribution of this, and it is:

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

So we will use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2.$$

To show that this can serve as a Pivot, and to use it, we need to find the Distribution of Y. Fortunately, someone already found the Distribution of this, and it is:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$

$\chi^2(n)$ Distribution

Definition: Assume

$$Z_1, Z_2, ..., Z_n \sim \mathcal{N}(0,1)$$

and Z_k-s are Independent (so they are IID). The Distribution of

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2$$

is called the **Chi-Squared Distribution with n Degrees of Freedom**, and is denoted by $\chi^2(n)$:

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2 \sim \chi^2(n).$$

$\chi^2(n)$ Distribution

Definition: Assume

$$Z_1, Z_2, ..., Z_n \sim \mathcal{N}(0,1)$$

and Z_k-s are Independent (so they are IID). The Distribution of

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2$$

is called the **Chi-Squared Distribution with n Degrees of Freedom**, and is denoted by $\chi^2(n)$:

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2 \sim \chi^2(n).$$

Say, if $Z \sim \mathcal{N}(0,1)$, then

$$Z^2 \sim$$

$\chi^2(n)$ Distribution

Definition: Assume

$$Z_1, Z_2, ..., Z_n \sim \mathcal{N}(0,1)$$

and Z_k-s are Independent (so they are IID). The Distribution of

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2$$

is called the **Chi-Squared Distribution with n Degrees of Freedom**, and is denoted by $\chi^2(n)$:

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2 \sim \chi^2(n).$$

Say, if $Z \sim \mathcal{N}(0,1)$, then

$$Z^2 \sim \chi^2(1).$$

$\chi^2(n)$ Distribution

Definition: Assume

$$Z_1, Z_2, ..., Z_n \sim \mathcal{N}(0,1)$$

and Z_k-s are Independent (so they are IID). The Distribution of

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2$$

is called the **Chi-Squared Distribution with n Degrees of Freedom**, and is denoted by $\chi^2(n)$:

$$Y = Z_1^2 + Z_2^2 + ... + Z_n^2 \sim \chi^2(n).$$

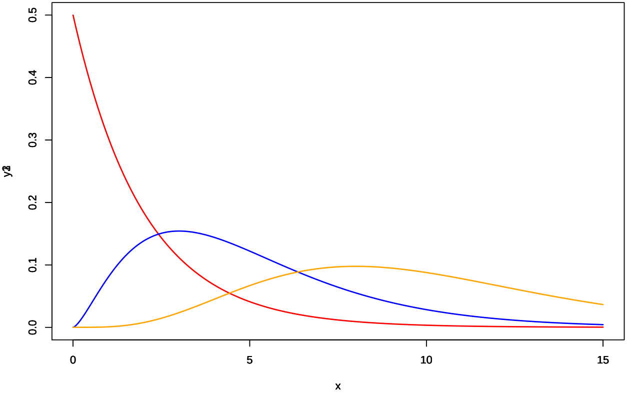
Say, if $Z \sim \mathcal{N}(0,1)$, then

$$Z^2 \sim \chi^2(1).$$

Exercise: What is $\mathbb{E}(Y)$?

$\chi^2(n)$ Distribution, PDF graphs

```
x <- seq(from = 0, to = 15, by = 0.01)
y1<-dchisq(x, df=2); y2<-dchisq(x, df=5); y3<-dchisq(x, df=10)
plot(x,y1,type="l",lwd=2,col="red", ylim=c(0,0.5)); par(new=T)
plot(x,y2,type="l",lwd=2,col="blue", ylim=c(0,0.5)); par(new=T)
plot(x,y3,type="l",lwd=2,col="orange", ylim = c(0,0.5))
```



CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

and we know now that

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

and we know now that

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

Now, as the general idea of Pivot CI construction was suggesting, we want to find a, b such that

$$\mathbb{P}(a < Y < b) = 1 - \alpha.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

and we know now that

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

Now, as the general idea of Pivot CI construction was suggesting, we want to find a, b such that

$$\mathbb{P}(a < Y < b) = 1 - \alpha.$$

The Distribution of Y, $\chi^2(n)$, is not symmetric, so it is not so easy to find the *minimum length CI*.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

and we know now that

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

Now, as the general idea of Pivot CI construction was suggesting, we want to find a, b such that

$$\mathbb{P}(a < Y < b) = 1 - \alpha.$$

The Distribution of Y, $\chi^2(n)$, is not symmetric, so it is not so easy to find the *minimum length CI*. And usually statisticians take a and b with

$$\mathbb{P}(Y \leq a) = \mathbb{P}(Y \geq b) = \frac{\alpha}{2},$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

Back to our CI for σ^2 Problem: We wanted to use the following as a Pivot:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2,$$

and we know now that

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

Now, as the general idea of Pivot CI construction was suggesting, we want to find a, b such that

$$\mathbb{P}(a < Y < b) = 1 - \alpha.$$

The Distribution of Y, $\chi^2(n)$, is not symmetric, so it is not so easy to find the *minimum length CI*. And usually statisticians take a and b with

$$\mathbb{P}(Y \leq a) = \mathbb{P}(Y \geq b) = \frac{\alpha}{2}, \quad i.e.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a =$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad \textit{and} \qquad b =$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad and \qquad b = \chi^2_{n,1-\frac{\alpha}{2}}.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad \textit{and} \qquad b = \chi^2_{n,1-\frac{\alpha}{2}}.$$

Now, the rest is easy: we obtained

$$\mathbb{P}(\chi^2_{n,\frac{\alpha}{2}} < Y < \chi^2_{n,1-\frac{\alpha}{2}}) = 1 - \alpha.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad \textit{and} \qquad b = \chi^2_{n,1-\frac{\alpha}{2}}.$$

Now, the rest is easy: we obtained

$$\mathbb{P}(\chi^2_{n,\frac{\alpha}{2}} < Y < \chi^2_{n,1-\frac{\alpha}{2}}) = 1 - \alpha.$$

We plug the value of Y:

$$\mathbb{P}\left(\chi^2_{n,\frac{\alpha}{2}} < \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma}\right)^2 < \chi^2_{n,1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad \textit{and} \qquad b = \chi^2_{n,1-\frac{\alpha}{2}}.$$

Now, the rest is easy: we obtained

$$\mathbb{P}(\chi^2_{n,\frac{\alpha}{2}} < Y < \chi^2_{n,1-\frac{\alpha}{2}}) = 1 - \alpha.$$

We plug the value of Y:

$$\mathbb{P}\left(\chi^2_{n,\frac{\alpha}{2}} < \sum_{k=1}^{n}\left(\frac{X_k - \mu}{\sigma}\right)^2 < \chi^2_{n,1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

and solve for σ^2:

$$\mathbb{P}\left(\frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,\frac{\alpha}{2}}}\right) = 1 - \alpha$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **known**

$$a = \chi^2_{n,\frac{\alpha}{2}} \qquad \textit{and} \qquad b = \chi^2_{n,1-\frac{\alpha}{2}}.$$

Now, the rest is easy: we obtained

$$\mathbb{P}(\chi^2_{n,\frac{\alpha}{2}} < Y < \chi^2_{n,1-\frac{\alpha}{2}}) = 1 - \alpha.$$

We plug the value of Y:

$$\mathbb{P}\left(\chi^2_{n,\frac{\alpha}{2}} < \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma}\right)^2 < \chi^2_{n,1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

and solve for σ^2:

$$\mathbb{P}\left(\frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,\frac{\alpha}{2}}}\right) = 1 - \alpha$$

This means that we found a $(1-\alpha)$-level CI for σ^2:

$$\left(\frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \quad \frac{\sum_{k=1}^{n}(X_k - \mu)^2}{\chi^2_{n,\frac{\alpha}{2}}}\right).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

Fortunately, the story almost repeats itself: the Pivot for the previous case was:

$$Y =$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ. We take an α, and want to construct a CI of CL $1 - \alpha$ for σ^2, using a Pivotal Quantity.

Fortunately, the story almost repeats itself: the Pivot for the previous case was:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2.$$

Well, we cannot use this now, since μ is unknown to us. So we will adjust Y a little bit.

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

Fortunately, the story almost repeats itself: the Pivot for the previous case was:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2.$$

Well, we cannot use this now, since μ is unknown to us. So we will adjust Y a little bit. Can you give some suggestion?

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Problem: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{N}(\mu, \sigma^2),$$

where μ is **unknown**, but we are not interested in μ. We take an α, and want to construct a CI of CL $1-\alpha$ for σ^2, using a Pivotal Quantity.

Fortunately, the story almost repeats itself: the Pivot for the previous case was:

$$Y = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma} \right)^2.$$

Well, we cannot use this now, since μ is unknown to us. So we will adjust Y a little bit. Can you give some suggestion? Of course, thanks:

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2.$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Great thing is that this is a Pivot: the Distribution of

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2$$

is independent of σ^2, and is:

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Great thing is that this is a Pivot: the Distribution of

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2$$

is independent of σ^2, and is: (ta-da-da-daaam)

$$\chi^2 \sim$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Great thing is that this is a Pivot: the Distribution of

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2$$

is independent of σ^2, and is: (ta-da-da-daaam)

$$\chi^2 \sim \chi^2(n-1).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Great thing is that this is a Pivot: the Distribution of

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2$$

is independent of σ^2, and is: (ta-da-da-daaam)

$$\chi^2 \sim \chi^2(n-1).$$

Simple and beautiful $\ddot\smile$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Great thing is that this is a Pivot: the Distribution of

$$\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma} \right)^2$$

is independent of σ^2, and is: (ta-da-da-daaam)

$$\chi^2 \sim \chi^2(n-1).$$

Simple and beautiful $\ddot\smile$

Now, doing the same calculations as above, we will arrive at the following $(1-\alpha)$-level CI for σ^2 :

$$\left(\frac{\sum_{k=1}^{n} (X_k - \overline{X})^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{\sum_{k=1}^{n} (X_k - \overline{X})^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Again, we have obtained the following $(1-\alpha)$-level CI for σ^2 :

$$\left(\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right).$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Again, we have obtained the following $(1-\alpha)$-level CI for σ^2 :

$$\left(\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right).$$

Usually, you will see this in the following form:

$$\left(\frac{(n-1)\cdot S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{(n-1)\cdot S^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right),$$

CI for σ^2, $\mathcal{N}(\mu, \sigma^2)$ Model, μ is **unknown**

Again, we have obtained the following $(1-\alpha)$-level CI for σ^2 :

$$\left(\frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right).$$

Usually, you will see this in the following form:

$$\left(\frac{(n-1)\cdot S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \quad \frac{(n-1)\cdot S^2}{\chi^2_{n-1,\frac{\alpha}{2}}} \right),$$

where S is the Sample Standard Deviation given by

$$S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}.$$

Example

Example: Assume we have scale of very high precision, and we want to measure the precision of our scale. Say, we weight a 50AMD coin, using that scale, 10 times, and we obtain the following measurements (in gramms):

```
 [1] 3.449243 3.450802 3.453054 3.448778 3.452541 3.448835 3.448
 [8] 3.454406 3.450314 3.449047
```

Our aim is to Estimate the Precision of the Scale.