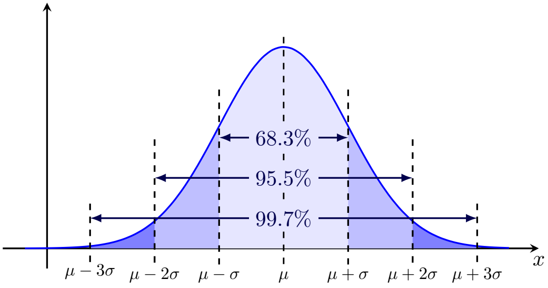
Confidence Intervals

EE-209 - Eléments de Statistiques pour les Data Sciences

The 68.3 - 95.5 - 99.7 rule

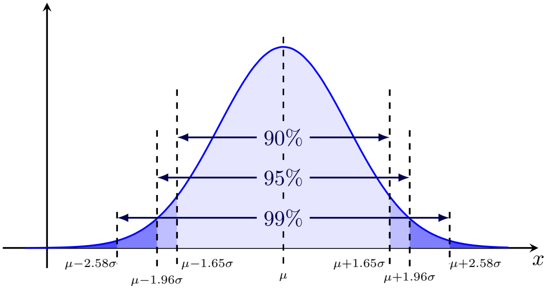


For $\quad X \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{P}(X \in [\mu - \sigma, \mu + \sigma]) \approx 0.683$$
$$\mathbb{P}(X \in [\mu - 2\,\sigma, \mu + 2\,\sigma]) \approx 0.955$$
$$\mathbb{P}(X \in [\mu - 3\,\sigma, \mu + 3\,\sigma]) \approx 0.997$$

We see that the probability that a Gaussian random variable takes a value which is further away from the expectation μ than 3σ (even 2 σ) is fairly small.

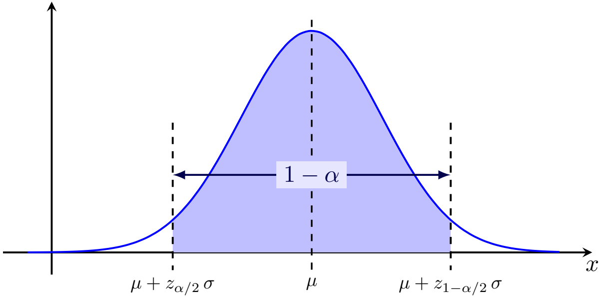
Intervals with guarantees at 90%, 95% and 99%



For $\quad X \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{P}(X \in [\mu - 1.645\,\sigma, \mu + 1.645\,\sigma]) \approx 0.90$$
$$\mathbb{P}(X \in [\mu - 1.960\,\sigma, \mu + 1.960\,\sigma]) \approx 0.95$$
$$\mathbb{P}(X \in [\mu - 2.576\,\sigma, \mu + 2.576\,\sigma]) \approx 0.99$$
$$\mathbb{P}(X \in [\mu - 3.291\,\sigma, \mu + 3.291\,\sigma]) \approx 0.999$$

High probability interval for a single Gaussian observation $X \sim \mathcal{N}(\mu, \sigma^2)$



$$
\begin{aligned}
& \mathbb{P}(\quad X \quad \in [\mu - q\,\sigma, \mu + q\,\sigma]) \\
= \; & \mathbb{P}(X - \mu \in [-q\,\sigma,\ q\,\sigma]) \\
= \; & \mathbb{P}\left(\frac{X-\mu}{\sigma} \in [-q,\ q]\right) \\
= \; & \mathbb{P}(\quad Z \quad \in [-q,\ q]),
\end{aligned}
$$

where $Z \sim \mathcal{N}(0,1)$.

And $\mathbb{P}(Z \in [-q, q]) = \Phi(q) - \Phi(-q) = 1 - 2\Phi(-q)$ where Φ is the standard Gaussian cdf.

So $\mathbb{P}(Z \in [-q, q]) = 1 - \alpha \quad \Leftrightarrow \quad -q = z_{\alpha/2} \quad \Leftrightarrow \quad q = z_{1-\alpha/2} = |z_{\alpha/2}|$.

We have $\qquad \mathbb{P}\big(X \in [\mu - z_{1-\alpha/2}\,\sigma, \mu + z_{1-\alpha/2}\,\sigma]\big) = 1 - \alpha$

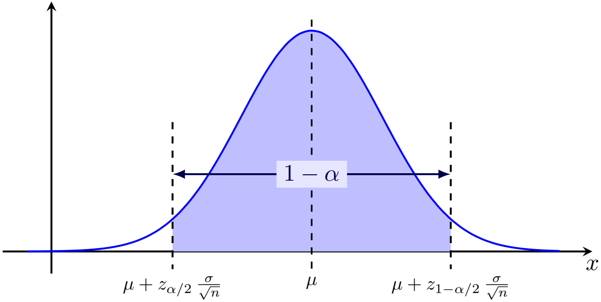
or equivalently $\qquad \mathbb{P}\big(X \in [\mu - |z_{\alpha/2}|\,\sigma, \mu + |z_{\alpha/2}|\,\sigma]\big) = 1 - \alpha$

High probability interval for the empirical mean of i.i.d. Gaussian data

$$\begin{array}{rccc} \text{If} & X_1, \ldots, X_n & \overset{\text{iid}}{\sim} & \mathcal{N}(\mu, \sigma^2) \\ \text{then} & X_1 + \ldots + X_n & \sim & \mathcal{N}(n\mu, n\sigma^2) \\ \text{so that} & \bar{X} & \sim & \mathcal{N}(\mu, \frac{\sigma^2}{n}) \end{array}$$

where $\bar{X} := \frac{1}{n}(X_1 + \ldots + X_n)$.

$\mathrm{std}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$ is called the *standard error*.



$$\begin{array}{rlcl} \text{So we have} & \mathbb{P}\Big(\bar{X} \in \big[\mu - z_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}, \mu + z_{1-\alpha/2}\,\frac{\sigma}{\sqrt{n}}\big]\Big) & = & 1-\alpha \\ \text{or equivalently} & \mathbb{P}\Big(\bar{X} \in \big[\mu - |z_{\alpha/2}|\,\frac{\sigma}{\sqrt{n}}, \mu + |z_{\alpha/2}|\,\frac{\sigma}{\sqrt{n}}\big]\Big) & = & 1-\alpha \\ \text{e.g.} & \mathbb{P}\Big(\bar{X} \in \big[\mu - 1.96\,\frac{\sigma}{\sqrt{n}}, \mu + 1.96\,\frac{\sigma}{\sqrt{n}}\big]\Big) & = & 0.95 \end{array}$$

Confidence interval: key idea

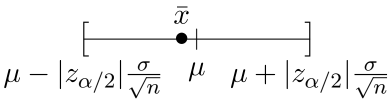
$$\mathbb{P}(\bar{X} \in [\mu - c, \mu + c]) = \mathbb{P}(|\bar{X} - \mu| \leq c) = \mathbb{P}(\mu \in [\bar{X} - c, \bar{X} + c])$$

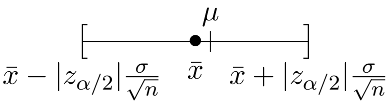
Indeed $\quad \bar{X} \leq \mu + c \quad \Leftrightarrow \quad \mu \geq \bar{X} - c. \qquad$ And $\quad \mu - c \leq \bar{X} \quad \Leftrightarrow \quad \mu \leq \bar{X} + c.$

So we have $\qquad \mathbb{P}\Big(\mu \in \Big[\bar{X} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}\Big]\Big) = 1 - \alpha$

or equivalently $\qquad \mathbb{P}\Big(\mu \in \Big[\bar{X} - |z_{\alpha/2}| \frac{\sigma}{\sqrt{n}}, \bar{X} + |z_{\alpha/2}| \frac{\sigma}{\sqrt{n}}\Big]\Big) = 1 - \alpha$

e.g. $\qquad \mathbb{P}\Big(\mu \in \Big[\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\Big]\Big) = 0.95$





- $\Big[\bar{X} - |z_{\alpha/2}| \frac{\sigma}{\sqrt{n}}, \bar{X} + |z_{\alpha/2}| \frac{\sigma}{\sqrt{n}}\Big]$ is a $1 - \alpha$ level *confidence interval.*
- $\Big[\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\Big]$ is a 95% level *confidence interval.*

It is the interval which is random, not μ !

Example: estimating a temperature

Let's assume that we are trying to measure a temperature (e.g. below a glacier) with a device which is fairly unstable. We assume that the standard deviation of the measurement error is known and equal to $\sigma = 0.6$.

We collect the following list of measured values:

$$[-0.1,\ 0.3, -0.8,\ 0.0,\ 0.1, -0.8,\ 0.7, -1.9, -0.9, -0.3]$$

We have $n = 10$ and $\bar{x} = -0.37$.
And we get the following 95% confidence interval:

$$\left[\bar{x} - 1.96\ \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\ \frac{\sigma}{\sqrt{n}}\right] = \left[-0.37 - 1.96 \cdot \frac{0.6}{\sqrt{10}}, -0.37 + 1.96 \cdot \frac{0.6}{\sqrt{10}}\right] = [-0.74, 0.00].$$

Note that in this example σ was known which is rarely the case

The case of the Gaussian empirical mean $\bar{X}$ when σ is unknown

We assume again that $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ so that $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$.

We can estimate σ^2 using the *unbiased variance estimate* $S^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$.

The way we proceeded before was using the fact that $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$...

But, given that S^2 is a random variable, a priori, we cannot simply replace σ^2 by S^2 in the previous equation.However we have the following result:

Theorem: the appropriately standardized mean follows a Student distribution...

Under the assumption that $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$:

(i) $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, (ii) $S^2/\sigma^2 \sim \frac{1}{n-1} \chi^2_{n-1}$, (iii) $\bar{X}$ and S^2 are independent r.v.s

and $T := \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ follows Student's t-distribution St_{n-1} with $n-1$ degrees of freedom.

The Student distribution

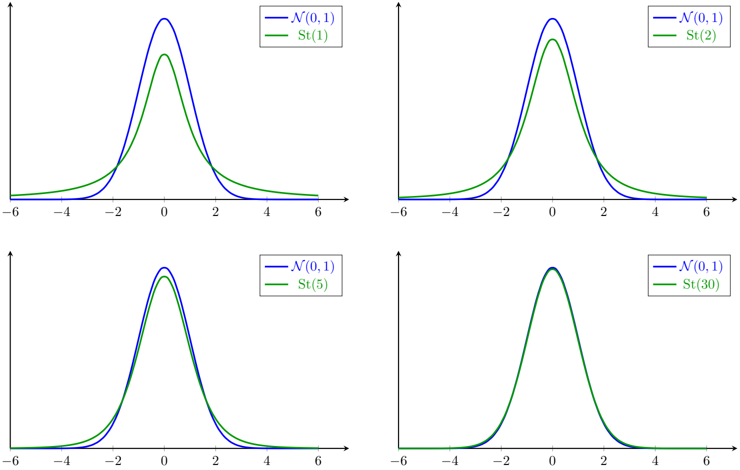
If $T \sim \mathrm{St}_n$ then its pdf is

$$p_T(t) = c_n \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}},$$

where c_n is a normalizing constant.

Away from the mean, the Gaussian density decreases *faster than exponentially*, while the Student t density decreases only *polynomially*.

Intervals containing 95% of the probability mass are thus wider for the Student.



Comparing the Student pdfs with $n = 1, 2, 5, 30$ d.f.s with a $\mathcal{N}(0,1)$.

Confidence Interval for the Gaussian $\bar{X}$ using the Student t-distribution

Given that $\quad T := \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim \mathrm{St}_{n-1}, \quad$ if

- $t^{(n-1)}_{\alpha/2}$ is the quantile of level $\frac{\alpha}{2}$ of a St_{n-1},
- $\tau_{\frac{\alpha}{2}} := t^{(n-1)}_{1-\alpha/2} = |t^{(n-1)}_{\alpha/2}|$ is the quantile of level $1 - \frac{\alpha}{2}$ of a St_{n-1},

then
$$
\begin{aligned}
1 - \alpha &= \mathbb{P}(-\tau_{\frac{\alpha}{2}} \leq T \leq \tau_{\frac{\alpha}{2}}) \\
&= \mathbb{P}(-\tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \leq \bar{X} - \mu \leq \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}) \\
&= \mathbb{P}(-\tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \leq \mu - \bar{X} \leq \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}) \\
&= \mathbb{P}(\bar{X} - \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})
\end{aligned}
$$

So $\left[\bar{X} - \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + \tau_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right]$ is a confidence interval of level $1 - \alpha$

Comparing $t^{(n)}_{1-\alpha/2}$ and $z_{1-\alpha/2}$

	Values of $\tau_{\alpha/2}$ for $n =$								$\lvert z_{\alpha/2}\rvert$
$1-\alpha$	1	2	5	10	20	50	100	∞	
0.90	6.31	2.92	2.02	1.81	1.73	1.67	1.66	1.645	1.645
0.95	12.7	4.30	2.57	2.23	2.09	2.01	1.98	1.960	1.960
0.99	63.6	9.92	4.03	3.17	2.85	2.68	2.62	2.576	2.576

- For $n = 6$, $\left[\bar{X} - 2.57\frac{S}{\sqrt{n}}, \bar{X} + 2.57\frac{S}{\sqrt{n}}\right]$ is a confidence interval of level 95% for μ.
- For $n = 51$, $\left[\bar{X} - 2.01\frac{S}{\sqrt{n}}, \bar{X} + 2.01\frac{S}{\sqrt{n}}\right]$ is a confidence interval of level 95% for μ.

When $n \to \infty$ we have $t^{(n)}_{1-\alpha/2} \longrightarrow z_{1-\alpha/2} = |z_{\alpha/2}|$.

Example: estimating a temperature again but with σ unknown

We are still trying to measure a temperature (e.g. below a glacier) with a device which is fairly unstable, but now σ is unknown.
We have the same list of measured values:

$$[-0.1,\ 0.3, -0.8,\ 0.0,\ 0.1, -0.8,\ 0.7, -1.9, -0.9, -0.3]$$

and we still have $n = 10$ and $\bar{x} = -0.37$.

We compute the sample standard deviation $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0.747$

And we get the following 95% Student confidence interval using that $t^{(9)}_{0.975} = 2.26$.

$$\left[\bar{x} - \tau_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \bar{x} + \tau_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right] = \left[-0.37 - 2.26 \cdot \frac{0.747}{\sqrt{10}},\ -0.37 + 2.26 \cdot \frac{0.747}{\sqrt{10}}\right] = [-0.90,\ 0.16].$$

The confidence interval is larger because $s > \sigma$ and because $t_{1-\alpha/2} > z_{1-\alpha/2}$.

Asymptotic Confidence Intervals

- What if $X_1, \ldots, X_n \overset{\text{iid}}{\sim} P$ with P unknown?
- Can we still determine a confidence interval for $\mu = \mathbb{E}[X_1]$?
- If we assume that $\mathbb{E}[X_1^2] < \infty$, then by the CLT, $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{(d)} \mathcal{N}(0,1)$.
- Even better, if $\mathbb{E}[X_1^2] < \infty$, by the CLT + Slutsky's lemma, $\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \xrightarrow{(d)} \mathcal{N}(0,1)$.

then if n is large

$$
\begin{aligned}
1 - \alpha \quad \approx \quad & \mathbb{P}\left(-|z_{\alpha/2}| \leq \tfrac{\bar{X}-\mu}{S/\sqrt{n}} \leq |z_{\alpha/2}|\right) \\
= & \, \mathbb{P}\left(-|z_{\alpha/2}|\tfrac{S}{\sqrt{n}} \leq \bar{X} - \mu \leq |z_{\alpha/2}|\tfrac{S}{\sqrt{n}}\right) \\
= & \, \mathbb{P}\left(-|z_{\alpha/2}|\tfrac{S}{\sqrt{n}} \leq \mu - \bar{X} \leq |z_{\alpha/2}|\tfrac{S}{\sqrt{n}}\right) \\
= & \, \mathbb{P}\left(\bar{X} - |z_{\alpha/2}|\tfrac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + |z_{\alpha/2}|\tfrac{S}{\sqrt{n}}\right)
\end{aligned}
$$

So $\left[\bar{X} - \tau_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \bar{X} + \tau_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$ is a *approximate* confidence interval of level $1 - \alpha$ when n is sufficiently large. This is called an asymptotic CI, and $1 - \alpha$ is its *nominal probability coverage*. Note that we used S here but any *consistent* estimator of σ could be used.

Asymptotic Confidence Intervals: Application to the Bernoulli

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Ber}(p)$ with p unknown, and we consider the estimator $\hat{p} := \bar{X}$.

- Given that $\mathbb{E}[X_1^2] = \mathbb{E}[X_1] = p < \infty$ and that $\text{Var}(X) = p(1-p)$, by the CLT

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{(d)} \mathcal{N}(0,1).$$

- $p(1-p)$ is unknown but can be estimated consistently by $\hat{p}(1-\hat{p})$, and by Slutsky

$$\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})}/\sqrt{n}} \xrightarrow{(d)} \mathcal{N}(0,1).$$

We therefore have the asymptotic CI at level $1-\alpha$ *(nominal probability coverage)*

$$p \in \left[\hat{p} - |z_{\alpha/2}|\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}, \hat{p} + |z_{\alpha/2}|\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\right].$$

Wald confidence intervals for the Maximum Likelihood Estimator

By the CLT, if $\hat{\theta} = \hat{\theta}_{\mathsf{MLE}}$ is the maximum likelihood estimator for θ based on an i.i.d. sample of size n, and if $I_1(\theta) > 0$, then

$$\sqrt{nI_1(\theta)}(\hat{\theta} - \theta) = \sqrt{I(\theta)}(\hat{\theta} - \theta) \xrightarrow{(d)} \mathcal{N}(0, 1).$$

With Slutsky's lemma, we also have $$\sqrt{I(\hat{\theta})}(\hat{\theta} - \theta) \xrightarrow{(d)} \mathcal{N}(0, 1).$$

So we have

$$\begin{aligned} 1 - \alpha \quad \approx \quad & \mathbb{P}(-|z_{\alpha/2}| \leq \sqrt{I(\hat{\theta})}(\hat{\theta} - \theta) \leq |z_{\alpha/2}|) \\ & = \mathbb{P}\left(\hat{\theta} - \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{\theta})}} \leq \theta \leq \hat{\theta} + \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{\theta})}}\right) \end{aligned}$$

Finally, $\left[\hat{\theta} - \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{\theta})}}, \; \hat{\theta} + \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{\theta})}}\right]$ is an asymptotic confidence interval of level $1 - \alpha$ which is thus valid when n is large.

Wald CI for the MLE of the parameter p in the Bernoulli model

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Ber}(p_*)$. If $N = \sum_{i=1}^n X_i = n\bar{X}$. then

- the log-likelihood is $\ell(p) = N \log p + (n - N) \log(1 - p)$.
- the score function is $\ell'(p) = \frac{N}{p} - \frac{n-N}{1-p} = \frac{(1-p)N - p(n-N)}{p(1-p)} = \frac{N-pn}{p(1-p)}$.
- the stationary points of ℓ satisfy $\ell'(p) = 0$. The unique solution is $\hat{p} = \frac{N}{n} = \bar{X}$.
- $\ell'(p) > 0$ for $p < \hat{p}$ and $\ell'(p) < 0$ for $p > \hat{p}$ so $\hat{p}$ attains the maximum and is the MLE.
- the Fisher Information is $I(p) = \text{Var}\big(\ell'(p)\big) = \frac{\text{Var}(N)}{p^2(1-p)^2} = \frac{np(1-p)}{p^2(1-p)^2} = \frac{n}{p(1-p)}$.
- It can be estimated by the observed information $I(\hat{p}) = \frac{n}{\hat{p}(1-\hat{p})}$.

Using the definition of the Wald confidence interval we have $p \in \left[\hat{p} - \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{p})}}, \hat{\theta} + \frac{|z_{\alpha/2}|}{\sqrt{I(\hat{p})}}\right]$

After replacement $p \in \left[\hat{p} - |z_{\alpha/2}| \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + |z_{\alpha/2}| \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$.This is the same asymptotic confidence interval as the one we had obtained from applying the general CLT.

Confidence Interval: general view

In the case of a scalar parameter θ

- instead of looking for a *pointwise estimator* $\hat{\theta}$ which aims at being close to θ
- we try to find a (short) interval $[\hat{\Theta}_l, \hat{\Theta}_u]$ such that

$$\mathbb{P}\big(\theta \in [\hat{\Theta}_l, \hat{\Theta}_u]\big) \geq 1 - \alpha.$$

- This interval is often of the form $[\hat{\theta} - m, \hat{\theta} + m]$ where m is the *margin of error* (MOE).
- Often, $m = q\frac{\sigma}{\sqrt{n}}$ or $m = q\frac{\hat{\sigma}}{\sqrt{n}}$ where q is the quantile of a distribution that does not depend on any (unknown) parameter, where σ is the standard deviation of a single observation and $\frac{\sigma}{\sqrt{n}}$ is called the *standard* error (SE).
- A confidence interval is a way to quantify our *uncertainty* about our estimate, and the MOE and SE are ways to measure it.
- When an approximate confidence interval is built it targets a level $1 - \alpha$, which is called the *nominal probability coverage*.
- It can be different from the actual value of $\mathbb{P}\big(\theta \in [\hat{\Theta}_l, \hat{\Theta}_u]\big)$, which is called the *actual probability coverage* and which is often unknown.

Pivots and how to construct confidence intervals

A **pivot** is a statistic $T(X_1, \ldots, X_n, \theta)$ that depends on the sample and on the parameter of interest in such a way that its distribution does not depend on θ.

For example:

- For $X_i \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\quad T := \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ is a pivot.
- For $X_i \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\quad T := \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \mathrm{St}_{n-1}$ is a pivot, for $S^2 = \frac{n}{n-1}(\overline{X^2} - \bar{X}^2)$.

We can also have some asymptotic pivots:

- For $X_i \ldots, X_n$ i.i.d. with $\mathbb{E}[X_1^2] < \infty$, $\quad T := \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} \xrightarrow{(d)} \mathcal{N}(0,1)$.
- For $\hat{\theta} = \hat{\theta}_{MLE}$, $\quad T := \sqrt{I(\hat{\theta})}(\hat{\theta} - \theta) \xrightarrow{(d)} \mathcal{N}(0,1)$.

Wilson score CI for the MLE of the parameter p in the Bernoulli model

We can also consider the central limit theorem based on the true variance $p(1-p)$:

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \xrightarrow{(d)} \mathcal{N}(0,1) \quad \text{so that} \quad \frac{n(\hat{p}-p)^2}{p(1-p)} \xrightarrow{(d)} \chi^2_1.$$

Let $z = z_{1-\alpha/2}$ be an $1-\alpha/2$ normal quantile. Then z^2 is an $1-\alpha$ quantile of the χ^2_1.

Therefore

$$\mathbb{P}\Big(\frac{n(\hat{p}-p)^2}{p(1-p)} \leq z^2\Big) \xrightarrow[n\to\infty]{} 1-\alpha.$$

The inequality can be rewritten as

$$np^2 - 2np\hat{p} + n\hat{p}^2 \leq z^2 p(1-p)$$

$$p^2(n+z^2) - 2p(n\hat{p} + \tfrac{1}{2}z^2) + n\hat{p}^2 \leq 0$$

Calculations show that this is equivalent to

$$p \in \Big[\hat{p}_z - \frac{\hat{\sigma}_z}{\sqrt{n}}\, z,\ \hat{p}_z + \frac{\hat{\sigma}_z}{\sqrt{n}}\, z\Big]$$

with $\hat{p}_z := \dfrac{n\hat{p} + z^2\frac{1}{2}}{n+z^2}$,

and $\hat{\sigma}_z := \dfrac{n}{n+z^2}\sqrt{\hat{p}(1-\hat{p}) + \dfrac{z^2}{4n}}$.