Statistical inference: Point estimation and Confidence Intervals
Introduction to statistical inference

From this chaper, we begin to learn how to make inference about unknown distribution given an observed data.

Parameter Point Estimation

In statistical analysis, we usually impose a distributional assumption on the random sample, that is, the sample's distribution belongs to certain family of distribution, say normal or gamma. Under such assumption, the unknown distribution is equivalent to the unknown parameter, therefore the inference problem reduces to the problem of estimating these parameters.

Definition:
A **point estimator** $\hat{\theta}$ of a parameter θ is the statistic used to estimate parameter from the sample. A **point estimate** is the value of point estimator given a specific sample.

Note:
- Point estimator is random, and point estimate is fixed single value.
- Literally, any statistic can be used as a point estimate. It is just a matter of whether it is a good or bad estimate.

Example
Sensible estimations for the population mean include:
- sample mean
- sample mean of sub-sample
- sample median, if the distribution is known to be symmetric
- If the distribution is gamma, we can estimate mean by $\hat{\alpha}\hat{\beta}$, where $\hat{\alpha}$ and $\hat{\beta}$ are two estimators of α and β .

Question: which one is the best?

Evaluation of point estimator
The point estimate value is a sampled value from the sampling distribution of estimator $\hat{\theta}$.
A good estimator should have a sampling distribution that closely centers around true parameter, this ensure that point estimate value based on single data set is close to the true parameter.

Bias of point estimator: $E\hat{\theta}-\theta$.
An estimator is called **unbiased** if its bias is 0. Unbiasness mean that **on average** the estimate we obtained is close to true parameter.

Example
Sample mean is an unbiased estimator for population mean.

Sample variance is an unbiased estimator for population variance.

Sample standard deviation is not unbiased.

$X_1,\ldots,X_n$ are independent random variable of Unif(0, θ). Denote estimator $\hat{\theta}=max\, X_i$. Is it unbiased?

<u>Variance of point estimator</u>
Given several unbiased estimators, we use variance to further compare them. Smaller variance means stabler estimation, means smaller variation between sample to sample.
<u>Standard error:</u> $\sqrt{Var(\hat{\theta})}$

Example:
$X_1,\ldots,X_n$ are independent random variable of Unif(0, θ). Denote two estimators

$$\hat{\theta}_1=\frac{n+1}{n}\, max\, X_i, \text{ and } \hat{\theta}_2=2\bar{X}$$

Compare these two estimators.

Theorem: Sample mean is the best estimator for the population mean if the data follows normal distribution.

Estimate the Variance of point estimator

One usually need to report the variance of estimator as reference for comparison.

Example:

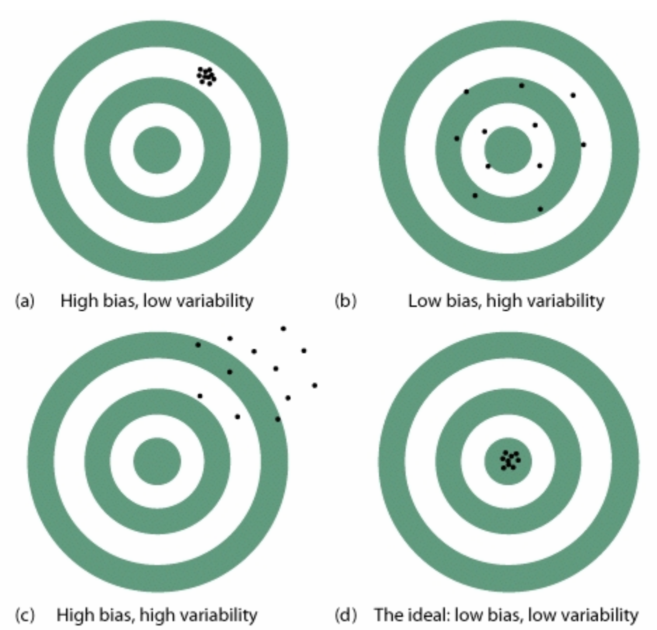
$X_1,\ldots,X_n$ are independent random variables of Normal(μ,σ^2). The variance of sample mean is σ^2/n , but σ^2 is unknown.

Estimated Standard Error: $\sqrt{s^2/n}$, that is to replace the unknown variance by the estimated variance.

Generally, if the standard error contains unknown parameter, we can simply replace it by the estimate of this parameter.

Mean squared error: Trade-off between bias and variance.

$$MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2=(E(\hat{\theta})-\theta)^2+Var(\hat{\theta}).$$



Parameter Estimation Methods

Methods of Moment

k-th (theoretical) moment of variable X is defined as

$$EX^k = \int x^k f(x;\theta)\,dx = \text{some function of } \theta .$$

k-th sample moment given a random sample is defined as $\left(\sum_{i=1}^n X_i^k\right)/n$.

Definition: moment estimator is obtained by solving estimation equations

$$\left(\sum_{i=1}^n X_i^k\right)/n = EX^k,$$

for k=1,2,….

Example:

1) $X_1,\ldots,X_n$ are Normal distribution with unknown mean and variance:

The likelihood function is

$$\left(\sum_{i=1}^n X_i\right)/n = EX = \mu$$

$$\left(\sum_{i=1}^n X_i^2\right)/n = EX^2 = (EX)^2 + Var(X) = \mu^2 + \sigma^2$$

Therefore,

$\hat{\mu} =$

$\hat{\sigma}^2 =$

2) $X_1,\ldots,X_n$ are gamma distribution with unknown α and β :

$$\left(\sum_{i=1}^n X_i\right)/n = EX =$$

$$\left(\sum_{i=1}^n X_i^2\right)/n = EX^2 = (EX)^2 + Var(X) =$$

Therefore,

$\hat{\alpha} =$

$\hat{\beta} =$

Moment estimator may yield invalid estimate, e.g. negative σ^2 , non-integer value for binomial relative distribution, etc.

Example: refer to textbook

Moment estimator is not invariant to data transformation.

Example: $X_1,\ldots,X_n$ are Normal distribution with unknown mean and variance 1, the moment estimator for mean is sample mean $\mu=\sum(X_i)/n$. On the other side, $Y_1=\exp(X_1),\ldots,Y_n=\exp(X_n)$ follows lognormal distribution, the moment estimator is $\sum\exp(X_i)/n=\exp(\mu)$, and $\mu=\log(\sum\exp(X_i)/n)$.

Maximum Likelihood estimator

Likelihood function is of the form of joint distribution of the data $X_1,\ldots,X_n$, where we treat the parameter θ as the argument of the function, and treat the $X_1,\ldots,X_n$ as fixed value, i.e.,

$$l(\theta;x_1,\ldots,x_n)=f(x_1,\ldots,x_n;\theta).$$

Maximum likelihood estimator (MLE) is defined as the **valid** value of θ that maximize likelihood function. That is

$$\hat{\theta}=\operatorname{argmax} l(\theta;x_1,\ldots,x_n).$$

Example:

1) $X_1,\ldots,X_n$ are Unif(0, θ):

The likelihood function is

$$l(\theta;X_{1,}\ldots,X_n)=\frac{1}{\theta^n}\text{ if } X_1,\ldots,X_n\leqslant\theta\text{, and } l(\theta;X_{1,}\ldots,X_n)=0\text{ otherwise.}$$

Thus, $\hat{\theta}=$

2) $X_1,\ldots,X_n$ are Normal distribution with unknown mean and variance:

The likelihood function is

$$l(\mu,\sigma^2;X_{1,}\ldots,X_n)=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left(\frac{-\sum_i(X_i-\mu)^2}{2\sigma^2}\right)\text{, and}$$

$$-\log l(\mu,\sigma^2;X_{1,}\ldots,X_n)=\frac{\sum_i(X_i-\mu)^2}{2\sigma^2}+\frac{n}{2}\log\sigma^2=\frac{\sum_i(X_i-\bar{X})^2+n(\bar{X}-\mu)^2}{2\sigma^2}+\frac{n}{2}\log\sigma^2,$$

Therefore,

$\hat{\mu}=$

$\hat{\sigma}^2 =$

3) $X_1, \ldots, X_n$ are Binomial distribution with unknown n

MLE follow invariant principle:

$$\text{MLE for } h(\theta) = h(\text{MLE of } \theta).$$

Example: $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$.

Theorem (Importance of MLE)
Generally, MLE is almost unbiased, with nearly the smallest variance, especially when n is sufficiently large.

Interval Estimate: Confidence intervals

A point estimate provides no information about the precision of the estimation. Besides, in many cases, $P(\text{point estimator = parameter}) = 0,$ and the point estimate says nothing about how close it might be to the true parameter.

an interval estimate or confidence interval (CI) reports an entire interval of plausible values instead of a single point estimation.

A Motivating example

Let us consider estimating the unknown population mean of normal distribution where the population variance is known.

The best point estimator is $\bar{X}$. Now we want to incorporate the precision of $\bar{X}$ (that is information of sampling distribution) into the point estimator.

1) $\bar{X} \sim N(\mu, \sigma^2/n)$

2) $Pr(\bar{X} \in (\mu - Z_{\alpha/2}\sigma/\sqrt{n}, \mu + Z_{\alpha/2}\sigma/\sqrt{n})) = 1 - \alpha$

3) $Pr(\mu \in (\bar{X} - Z_{\alpha/2}\sigma/\sqrt{n}, \bar{X} + Z_{\alpha/2}\sigma/\sqrt{n})) = 1 - \alpha$

Definition: the (1-alpha) confidence interval for population mean under normal distribution is

$$(\bar{X} - Z_{\alpha/2}\sigma/\sqrt{n}, \bar{X} + Z_{\alpha/2}\sigma/\sqrt{n})$$

where $Z_{\alpha/2}\sigma/\sqrt{n}$ is called margin of error.

Interpreting this interval

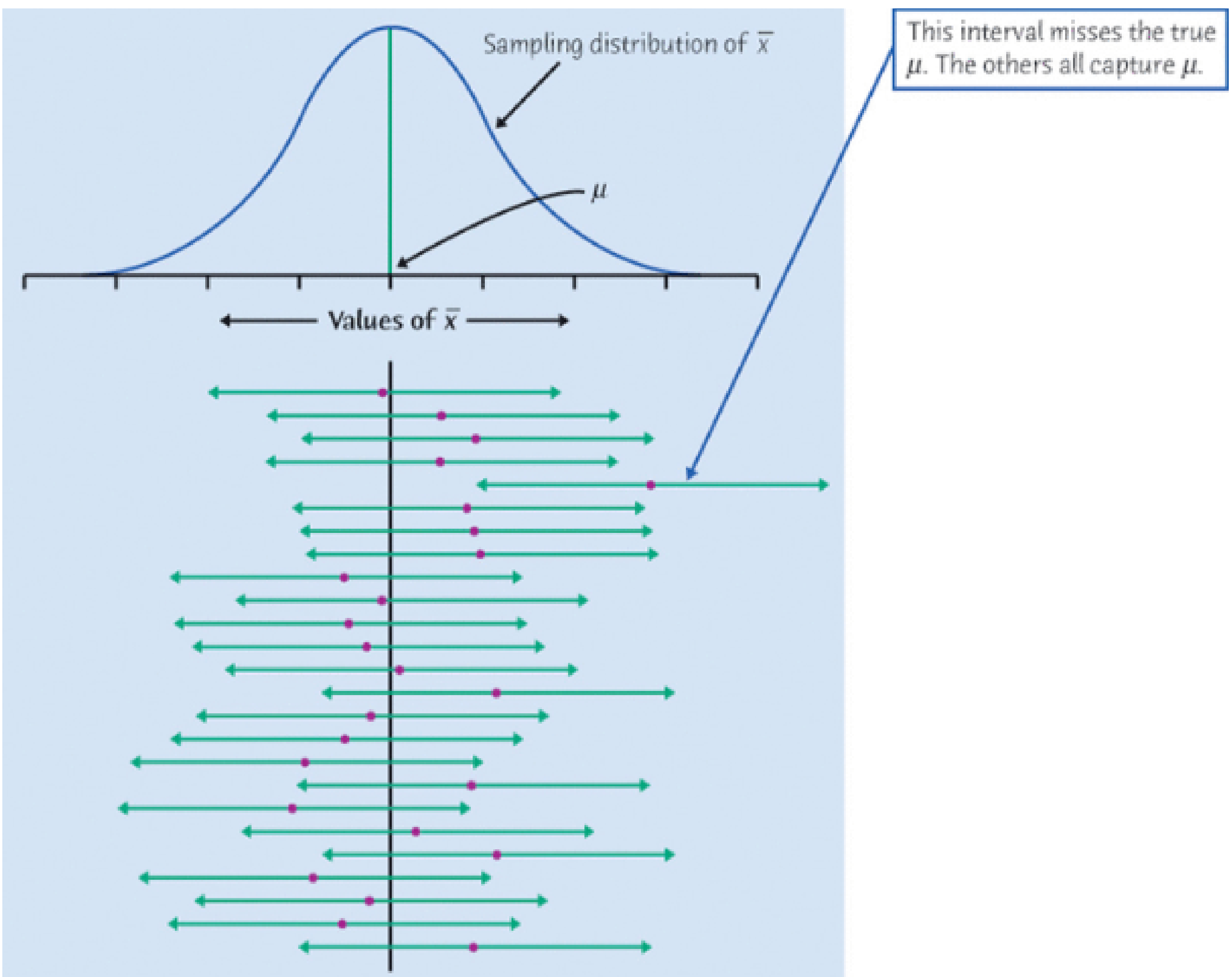
1. Confident statement: Given a fixed sample, we say that **with 1-alpha confidence**, the true parameter is inside interval $(\bar{x} - Z_{\alpha/2}\sigma/\sqrt{n}, \bar{x} + Z_{\alpha/2}\sigma/\sqrt{n})$.

   Why do we use the word ***confident?***

   ANSWER: Because we used a method that 95% of the time gives an interval that contains μ .

2. Probabilistic statement: **With probability 1-alpha**, random interval $(\bar{X} - Z_{\alpha/2}\sigma/\sqrt{n}, \bar{X} + Z_{\alpha/2}\sigma/\sqrt{n})$ contains the true parameter.

   Here, the probability are with respect to the randomness of the n-size sample. Given difference random sample, one can calculate different intervals. Some of them can cover the true parameter, some of them can not. And the probability you will randomly select a sample whose confidence interval *captures* the true value of μ is 95%.



R code for simulation of the coverage:

```
x <- matrix(rnorm(100000),ncol=10000)
mn <- apply(x,2,mean); v <- apply(x,2,var)
hwd <- 1.96/sqrt(10); lcl <- mn-hwd; ucl <- mn+hwd mean((lcl<0)&(ucl>0))   ## z-interval

hwd<-qt(.975,9)*sqrt(v/10); lcl<-mn-hwd; ucl<-mn+hwd
mean((lcl<0)&(ucl>0)) ## t-interval

x <- matrix(runif(100000),ncol=10000)
mn <- apply(x,2,mean); v <- apply(x,2,var)
hwd<-1.96*sqrt(1/120); lcl<-mn-hwd; ucl<-mn+hwd
mean((lcl<.5)&(ucl>.5)) ## z-interval
hwd<-qt(.975,9)*sqrt(v/10); lcl<-mn-hwd; ucl<-mn+hwd
mean((lcl<.5)&(ucl>.5)) ## t-interval
```

Example: Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under conditions of stress. A researcher conducted a study to investigate whether a program of regular exercise might affect the resting (unstressed) concentration of HBE in the blood. Ten adults were subjects in the study. Each subject's HBE level was measured at the beginning of the study. Then, each subject participated in a 5 month physical fitness program, at the end of which, each subject's HBE level was measured again. The average drop in HBE level was $\bar{x}=13.0$ (pg/mLi) The standard deviation of the change in HBE levels is known to be $\sigma=12.4$ (pg/mLi).

- What is the point estimate of the parameter?

- 95% <u>confidence interval</u> for μ:

- <u>Interpretation:</u> We are 95% confident that, for <u>all adults</u> participating in a 5 month physical fitness program, the average drop in HBE is

- Do you think there is statistical evidence the average HBE drop is *not* 18?

Key Question: Why can't we interpret our 95% confidence interval as: *There is a 95% probability that the average drop in HBE is between 5.3 and 20.7* (pg/mLi) ?

<u>Key point:</u> A confidence interval is an inference about the population based on a calculation using data from a sample. A confidence interval only makes an inference about a population parameter. CI's don't tell us anything about individuals in the population. A confidence interval uses the sample mean in the calculation, but it is not an inference about the sample mean. We know all about the sample because we have data from the data. There is no reason to make inferences about the sample – either the subjects in the sample or statistics calculated from the sample.

Sample size, confidence level and precision

Proposition:

1. Larger the confidence level is, smaller the alpha is, and wider the interval is.

Confidence level C	90%	95%	99%
Critical value z^*	1.645	1.960	2.576

2. Larger the sample size is, narrower the interval is.

3. If we want to control the width to be smaller than w, (that is margin of error is smaller than w/2), the necessary sample size is at least $\left(2 Z_{\alpha/2} \sigma / w\right)^2$

A formal, general definition of confidence interval:

A confidence interval consists of **two statistics** $L(X_1,\ldots,X_n)$ and $U(X_1,\ldots,X_n)$, such that

$$P\left(L(X_1,\ldots,X_n) < \theta < U(X_1,\ldots,X_n)\right) = 1-\alpha.$$

Note that one of L and U can be infinite, this lead to **Confidence Bounds:** for example, for normal mean problem,

$$Pr\left(\mu \in \left(-\infty, \bar{X} + Z_\alpha \sigma/\sqrt{n}\right)\right) = 1-\alpha.$$

We can derive a confidence interval, starting for a point estimator and its associated sampling distribution.

Refer to Example 7.5 in textbook.

Large sample CI for population mean

Our CI formula in last section requires 1) normality, 2) known standard deviation. Now we consider if these conditions don't hold.

Suppose $X_1,\ldots,X_n$ are random sample with unknown mean μ and unknown variance σ^2.
First of all, even they are not normal rv's, by CLT, we still have the large sample approximation that

$$\bar{X} \approx N\left(\mu, \sigma^2/n\right) \text{ and } \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \approx Z.$$

If furthermore, if we have a good estimation $\hat{\sigma}$ for σ, we can plug in this estimator, then we obtain

$$\frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} \approx \frac{\sigma}{\hat{\sigma}} \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \approx Z.$$

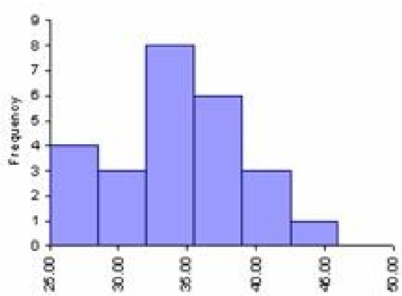
One popular estimator for population standard deviation is sample standard deviation. Thus, it leads to the following **large sample approximated confident interval** with 1-alpha confidence level:

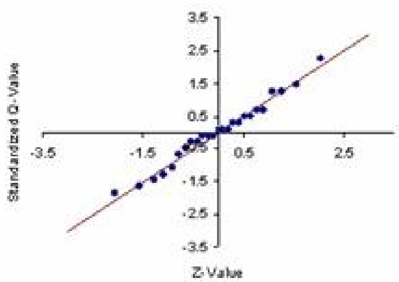
$$(\bar{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}}, \bar{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}})$$

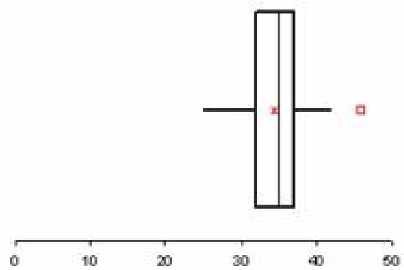
<u>Note</u> The margin of error here is a not fixed value, but data-dependent.

Example: The FFQ is a food frequency questionnaire, which uses a carefully designed series of questions to determine the dietary intakes of participants in the study. In the Nurses' Health Study, a random sample of 166 female nurses each completed the FFQ. The variable being measured is the % fat in the diet.

$\bar{x} =$ 36.50 and s = 6.75.







Find the approximated confidence intervals for μ = average% fat in the diet of all women similar to the female nurses in this study during the time period that this study was conducted

I. 90% CI: $\bar{x} \pm 1.645\frac{6.75}{\sqrt{166}} = 36.5 \pm 0.86$

II. What happens to the margin of error as the confidence level increases?

III. What is the correct interpretation of the 90% confidence interval we calculated?

<u>Confidence Intervals for Proportions</u>

Suppose $X_1, \ldots, X_n$ are random Bernoulli sample with unknown true success rate p.
Note that EX=p, Var(X)=p(1-p).

To estimate a population proportion *p*, we use the sample proportion $\hat{p} = \bar{X}$. Similarly, by CLT

$$\frac{\hat{p} - p}{\sqrt{Var(X)/n}} \approx Z.$$

There are several way to derive a valid CI from this point.

- By CLT approximation,

$$P\left(-z_{\alpha/2} \leqslant \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}\right) \approx 1-\alpha .$$

  The inequality inside the bracket can be transfer to an inequality of p, which leads to the score CI.

- We can estimate the variance by sample variance, this leads to CI
  $\left(\hat{p}-Z_{\alpha/2}\frac{S}{\sqrt{n}}, \hat{p}+Z_{\alpha/2}\frac{S}{\sqrt{n}}\right)$, where $S^2=n\,\hat{p}(1-\hat{p})/(n-1)$

- We estimate the unknown variance by $\hat{p}(1-\hat{p})$, this leads to CI
  $\left(\hat{p}-Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}, \hat{p}+Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right)$.

Small sample CI for population mean under normal distribution

Suppose $X_1,\ldots,X_n$ are normal random samples with unknown mean μ and unknown variance σ^2 .

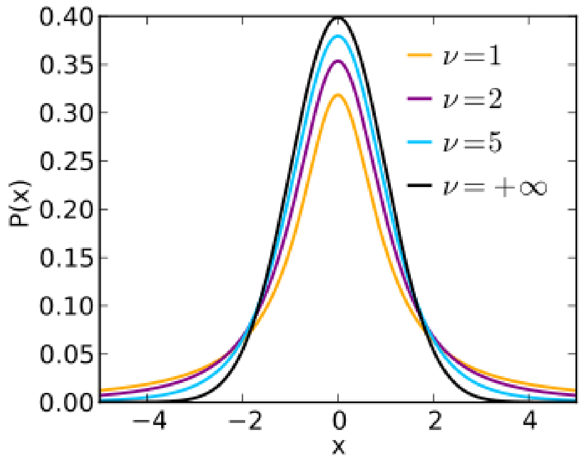
As we studied, $\frac{\bar{X}-\mu}{S/\sqrt{n}} \approx Z$ holds when sample size is sufficiently large. However, statisticians derive the exact distribution of $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ when all X's are normal.

Theorem: Under normality assumption, $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ has a continuous distribution call **student-t distribution** with n-1 degree of freedom (df).

Proposition of t distribution: Let t_v denote the t distribution with v df.

- It has a bell shape, and centered at 0.

- It is more spread out than normal.
- The df controls the variability of t distribution.



Notation: $t_{\alpha,v}$ denotes the t critical value satisfies

$$P(t_v > t_{\alpha,v}) = \alpha$$

Thus, $P\left(-t_{\alpha/2,v} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2,v}\right) = P(-t_{\alpha/2,v} < t_v < t_{\alpha/2,v}) = 1-\alpha$, and 1-alpha CI is $\bar{X} \pm t_{\alpha/2,v}\frac{S}{\sqrt{n}}$.

We can obtain the t critical value by R code qt(1-alpha, df=v), or use the T table:

T Table

				Confidence Level								
				80%	**90%**	**95%**		**98%**	**99%**		**99.8%**	**99.9%**
						Right-Tail Probability						
df	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
1	1.000	1.376	1.963	3.078	6.314	12.706	15.895	31.821	63.657	127.321	318.309	636.619
2	0.817	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.089	22.327	31.599
3	0.765	0.979	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.215	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.696	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965

18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
inf.	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090	3.291

If the *df* you want isn't in the table, use the *df* that is the closest smaller number.

Important:

The t confidence interval depends on normal assumption, one need to use histogram or QQ plot to diagnose whether this assumption is valid.
If the data presents strong non-normality, but the sample size is large enough, we can use large sample approximated CI.

Comparison between Z and t confidence interval: Same interval center, different interval width.
On average:

$$[(\text{margin of error for Z CI})^2] = Z^2_{\alpha/2}\sigma^2/n$$

$$[(\text{margin of error for t CI})^2] = E(t^2_{\alpha/2}S^2/n) = t^2_{\alpha/2}\sigma^2/n > Z^2_{\alpha/2}\sigma^2/n.$$

Intuitively, narrower CI means preciser interval estimation. For normal means estimation, if population standard deviation is known, then both z- and t- confidence interval are statistical valid CI.
Question: Can we compute both of them and pick the narrow one?

```
R code for simulation:
coverage.count<-0;
for(iter in 1:10000)
{  x<-rnorm(12);
me1<-qnorm(0.975)/sqrt(12);
me2←qt(0.975,11)*sd(x)/sqrt(12);
me<- min(me1,me2);
if(abs(mean(x))<=me) coverage.count<-coverage.count+1;
}
coverage.count/10000
```

Prediction Interval for a new observation from this normal distribution:

Let X_{n+1} be a new independent observation. We can show that

$$T=\frac{X_{n+1}-\bar{X}}{s\sqrt{1+\frac{1}{n}}}\sim t_{n-1}$$

Therefore, $Pr\left(X_{n+1}\in\bar{X}\pm t_{\alpha/2,n-1}S\sqrt{1+\frac{1}{n}}\right)=1-\alpha$.

Comparison between CI and PI:

CI is the interval estimation of E(X), PI is the interval estimation of X=(EX+randomness).
PI is always larger than CI, since PI also takes account of the variability of X.

Small sample CI for population variance under normal distribution

Now we are interested in providing a CI for the unknown population variance.

The point estimator is $S^2=\sum(X_i-\bar{X})^2/(n-1)$, to derive a CI, we want to figure out the distribution of sample variance.

Theorem: $\frac{(n-1)S^2}{\sigma^2}$ follows a chi-squred distribution with degree of freedom n-1.

Notation: $\chi^2_{\alpha,v}$ denotes the chi-squred critical value satisfies $P(\chi^2_v > \chi^2_{\alpha,v}) = \alpha$, thus

$$P(\chi^2_{1-\alpha_1,v} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha_2,v}) = P(\chi^2_{1-\alpha_1,v} < \chi^2_v < \chi^2_{\alpha_2,v}) = 1-\alpha_1-\alpha_2$$

This implies the following $1-\alpha_1-\alpha_2$ level CI:

$$\frac{(n-1)S^2}{\chi^2_{\alpha_2,v}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha_1,v}}.$$

Conventionally, we choose $\alpha_1 = \alpha_2 = \alpha/2$.

Note: This is not a symmetric confidence interval, but this confidence interval guarantees positiveness.

R code:

```
x<-rnorm(14)
t.test(x, conf.level=.95)

heads <- rbinom(1, size = 100, prob = .5)
prop.test(heads, 100, conf.level=.95)

n<-length(x);
var(x)*(n-1)/qchisq( c(0.975, 0.025), df=n-1 )
```