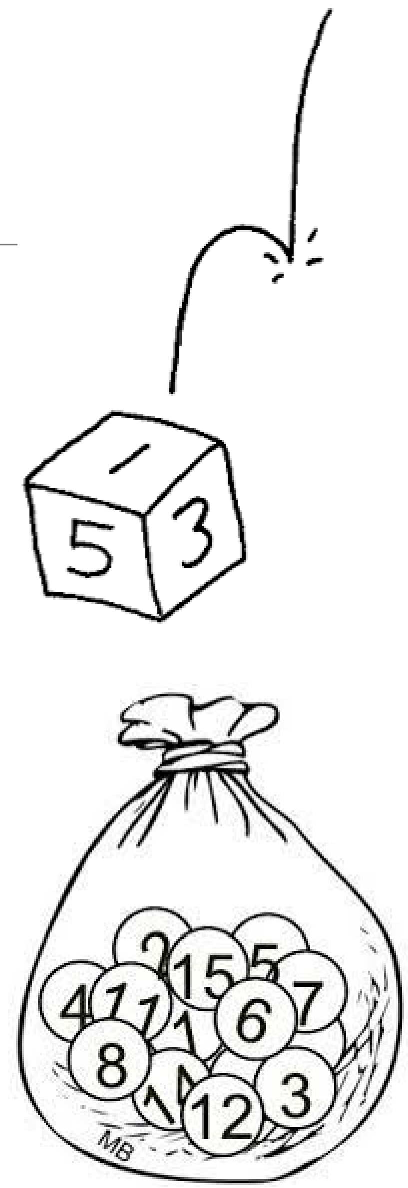
Lecture 17

- Moment Estimator
- Estimator Selection Criteria
- Confidence Intervals



Moment Estimator

矩估计

moments estimation is based solely on the law of large numbers.

---- math.arizona.edu

the method of moments involves equating sample moments with theoretical moments.

---- stat.psu.edu

Moment of data

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Consider a sample of n i.i.d. random variables $X_1, X_2, \dots, X_n$.

- X_i from distribution $F = B(p)$ with unknown parameter.
- Observed data:

$$[0, 0, 1, 1, 1, 1, 1, 1, 1, 1] \qquad (n = 10)$$

What is the moment of the observed data (样本矩)?

$$A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i^1 = \frac{1}{10}(0 + 0 + \cdots + 1 + 1) = 0.8$$

For Bernoulli distribution

$$\mu_1 = E(X^1) = p$$

Key: Finding connection between μ_1 and p

Thus,

$$\hat{p} = A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Moment of data

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Consider a sample of n i.i.d. random variables $X_1, X_2, \dots, X_n$.

- X_i from distribution $F = U(a,b)$ with unknown parameter a, b.
- Observed data:

$$[x_1, x_2, \dots, x_{10}] \qquad (n = 10)$$

What is the connection between μ_k and (a,b)?

1. $\mu_1 = E(X) = \frac{a+b}{2}$
2. $\mu_2 = E\left(X_i^2\right) = D(X_i) + \mu_1^2 = \frac{(b-a)^2}{12} + \frac{(a+b)^2}{4}$

Estimating the moments from data: $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

Solve a, b from A_k:

$$\begin{cases} \dfrac{a+b}{2} = A_1 \\ \dfrac{(b-a)^2}{12} + \dfrac{(a+b)^2}{4} = A_2 \end{cases}$$

$$\begin{cases} \hat{a} = A_1 - \sqrt{3(A_2 - (A_1)^2)} \\ \hat{b} = A_1 + \sqrt{3(A_2 - (A_1)^2)} \end{cases}$$

Moment estimator

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

def The moment estimator of θ finds:

the value(s) of θ that solves the moment equation(s)

Suppose we need to estimate $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ in $F(x; \theta_1, \theta_2, \dots, \theta_k)$, by one observed sample $X_1, X_2, \dots, X_n$.
Let μ_k denotes the k-th order raw moment.

$$\begin{cases} A_1 = E(X) = \mu_1(\theta_1, \theta_2, \dots, \theta_k) \\ A_2 = E(X^2) = \mu_2(\theta_1, \theta_2, \dots, \theta_k) \\ \quad\cdots\cdots\cdots\cdots\cdots\cdots \\ A_k = E(X^k) = \mu_k(\theta_1, \theta_2, \dots, \theta_k) \end{cases} \qquad \begin{cases} \hat{\theta}_1 = \theta_1(A_1, A_2, \dots, A_k) \\ \hat{\theta}_2 = \theta_2(A_1, A_2, \dots, A_k) \\ \quad\cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{\theta}_k = \theta_k(A_1, A_2, \dots, A_k) \end{cases}$$

$\hat{\theta} = \left(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k\right)$ is the moment estimator.

Steps in moment estimator

1. According to raw moment definition, establish
   $\mu_i = E\left[X^i\right] = \mu_i(\theta_1, \theta_2, \dots, \theta_k),\ i = 1,2, \dots, k$

2. Compute sample moments from the observations,
   $A_i = \mu_i(\theta_1, \theta_2, \dots, \theta_k),\ i = 1,2, \dots, k$

3. Formulate a group of equation,

$$\hat{\theta}_i = \theta_1(A_1, A_2, \dots, A_k),\ i = 1,2, \dots, k$$

Sample values →	Sample moments →	Raw moments →	Parameters
$X_1, X_2, \dots, X_n$	$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$	$E\left[X^i\right] = \mu_i(\theta_1, \theta_2, \dots, \theta_k)$	$\hat{\theta}_i = \theta_1(A_1, A_2, \dots, A_k)$

Data $\longrightarrow$ Parameters

Parameter estimation

- Point estimation
  - Moment estimator
  - Maximum Likelihood Estimator
  - other
- Interval estimation

Moment estimator

1. According to raw moment definition, establish $\mu_i = E\left[X^i\right] = \mu_i(\theta_1, \theta_2, \dots, \theta_k)$
2. Compute moments from the observed sample, $A_i = \mu_i(\theta_1, \theta_2, \dots, \theta_k)$
3. Formulate a group of equations and resolve $\hat{\theta}_i$, $\hat{\theta}_i = \theta_1(A_1, A_2, \dots, A_k)$

Statistics moments must exist.

Maximum Likelihood Estimator

1. Establish the likelihood function, $L(\theta)$:

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

2. Take natural log of the likelihood function:

$$LL(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(X_i|\theta)$$

3. Find the solution with derivative on $LL(\theta)$:

$$\frac{d \ln L(\theta)}{d\theta} = 0$$

$$\text{or} \quad \frac{\partial \ln L(\theta_i)}{\partial \theta_i} = 0, \ (i = 1,2,\dots,k)$$

1) know $f(X_i|\theta)$;
2) resolve the derivatives.

Ex. Given a population $X \sim \exp(\theta)$, where θ is unknown. $X_1, X_2, \ldots, X_n$ is a sample from X. Find the moment estimator of θ.

Sol.

We need one moment to estimate the single parameter θ.

$$\mu_1 = E(X) = \theta$$

where the sample moment can be estimated by

$$\mu_1 = A_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Therefore,

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

Ex. Given a population with $E(X) = \mu$, $D(X) = \sigma^2 \neq 0$, but μ, σ^2 are unknown. $X_1, X_2, \dots, X_n$ is a sample from X. Find the moment estimator of μ, σ^2. (unknown distribution!)

Sol. We need two equations to estimate two parameters.

$$\begin{cases} \mu_1 = E(X) = \mu \\ \mu_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2 \end{cases}$$

where the sample moment can be estimated by

$$A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i\,,\ A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Therefore,

$$\begin{cases} \hat{\mu} = A_1 = \bar{X} \\ \hat{\sigma}^2 = A_2 - A_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{cases}$$

Same μ, σ^2 moment estimators for different distributions.

Try by yourselves

Ex. Given a population follows

$$f(x) = \begin{cases} (\alpha+1)x^{\alpha}, 0 < x < 1 \\ 0, \qquad\qquad \text{otherwise} \end{cases}$$

where the value of α is unknown. $X_1, X_2, \ldots, X_n$ is a sample from X. Find the MLE and moment estimator of α.

Ex. Given the PMF of X as

X	1	2	3
p_i	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

where θ $(0 < \theta < 1)$ is unknown. Given the observed data

$$x_1 = 1, x_2 = 2, x_3 = 1$$

Find the MLE and moment estimator of θ.

Try by yourselves

Ex. Given a population follows

$$f(x) = \begin{cases} (\alpha+1)x^{\alpha}, 0 < x < 1 \\ 0, \qquad\qquad \text{otherwise} \end{cases}$$

where the value of α is unknown. $X_1, X_2, \ldots, X_n$ is a sample from X. Find the MLE and moment estimator of α.

Moment Estimator

$$E(X) = \int_0^1 x(\alpha+1)x^{\alpha}dx$$

$$= \frac{\alpha+1}{\alpha+2} = \bar{X}$$

$$\hat{\alpha} = \frac{2\bar{X}-1}{1-\bar{X}}$$

Maximum Likelihood Estimator

$L(\theta)$: $\quad L(\alpha) = \prod_{i=1}^{n}(\alpha+1)X_i^{\alpha}$

$$= (\alpha+1)^n \prod_{i=1}^{n} X_i^{\alpha}$$

$LL(\theta)$: $\quad \ln L(\alpha) = n \cdot \ln(\alpha+1) + \alpha \sum_{i=1}^{n} \ln X_i$

Derivative: $\quad \frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{n}{\alpha+1} + \sum_{i=1}^{n} \ln X_i = 0$

Find $\hat{\theta}$: $\quad \hat{\alpha} = -\frac{n}{\sum_{i=1}^{n} \ln X_i} - 1$

Try by yourselves

Ex. Given the PMF of X as

X	1	2	3
p_i	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

where θ $(0 < \theta < 1)$ is unknown. Given the observed data

$$x_1 = 1, x_2 = 2, x_3 = 1$$

Find the MLE and moment estimator of θ.

Moment Estimator

$$\begin{aligned} &E(X) \\ &= 1 \times \theta^2 + 2 \times 2\theta(1-\theta) \\ &\qquad + 3 \times (1-\theta)^2 \\ &= [\theta + 3(1-\theta)][\theta + (1-\theta)] \\ &= 3 - 2\theta = \bar{X}. \end{aligned}$$

$$\therefore \hat{\theta} = \frac{3-\bar{X}}{2} = \frac{3-\frac{1+2+1}{3}}{2} = \frac{5}{6}.$$

Maximum Likelihood Estimator:

Likelihood function:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{3} P\{X_i = x_i\} \\ &= P\{X_1 = 1\}P\{X_2 = 2\}P\{X_3 = 1\} \\ &= 2\theta^5(1-\theta). \end{aligned}$$

Take natural log:

$$\ln L(\theta) = \ln 2 + 5 \cdot \ln \theta + \ln(1-\theta).$$

Take derivative:

$$\frac{d \ln L(\theta)}{d\theta} = 0 + \frac{5}{\theta} - \frac{1}{1-\theta} = 0, \quad \therefore \ \hat{\theta} = \frac{5}{6}.$$

Data → Parameters

Parameter estimation

Point estimation | Interval estimation

Moment estimator | Maximum Likelihood Estimator | other

1. According to raw moment definition, establish $\mu_i = E[X^i] = \mu_i(\theta_1, \theta_2, \dots, \theta_k)$
2. Compute moments from the observed sample, $A_i = \mu_i(\theta_1, \theta_2, \dots, \theta_k)$
3. Formulate a group of equations and resolve $\hat{\theta}_i$, $\hat{\theta}_i = \theta_1(A_1, A_2, \dots, A_k)$

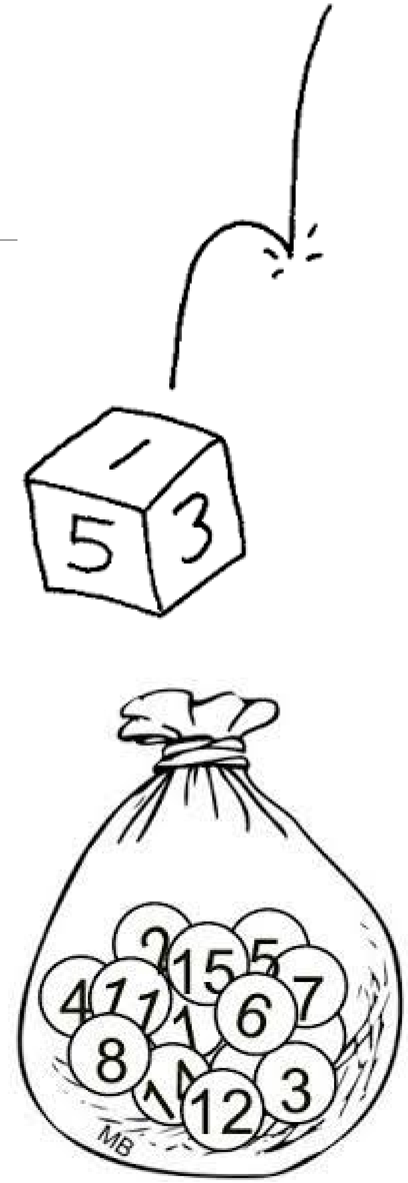
Statistics moments must exist.

1. Establish the likelihood function, $L(\theta)$:
$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$
2. Take natural $\log$ of the likelihood function:
$$LL(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(X_i|\theta)$$
3. Find the solution with derivative on $LL(\theta)$:
$$\frac{d \ln L(\theta)}{d\theta} = 0$$
$$\text{or} \quad \frac{\partial \ln L(\theta_i)}{\partial \theta_i} = 0, \ (i = 1,2,\dots,k)$$

1) know $f(X_i|\theta)$;
2) resolve the derivatives.

Lecture 17

- Moment Estimator
- Estimator Selection Criteria
- Confidence Intervals



Recall: Try by yourselves

Ex. Given a population follows

$$f(x) = \begin{cases} (\alpha + 1)x^{\alpha}, 0 < x < 1 \\ 0, \qquad\qquad \text{otherwise} \end{cases}$$

where the value of α is unknown. $X_1, X_2, \ldots, X_n$ is a sample from X. Find the MLE and moment estimator of α.

Moment estimator

$$E(X) = \int_0^1 x(\alpha + 1)x^{\alpha} dx$$

$$= \frac{\alpha+1}{\alpha+2} = A_1 = \bar{X}$$

$$\hat{\alpha} = \frac{2\bar{X}-1}{1-\bar{X}}$$

Maximum Likelihood Estimator

$$L(\alpha) = \prod_{i=1}^{n} (\alpha + 1)X_i^{\alpha} = (\alpha + 1)^n \prod_{i=1}^{n} X_i^{\alpha}$$

$$\ln L(\alpha) = n \cdot \ln(\alpha + 1) + \alpha \sum_{i=1}^{n} \ln X_i$$

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{n}{\alpha+1} + \sum_{i=1}^{n} ln\, X_i = 0$$

$$\hat{\alpha} = -\frac{n}{\sum_{i=1}^{n} ln\, X_i} - 1$$

Three important criteria

1. Biasedness (无偏性):
   Def. $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$
   if $\text{Bias}(\hat{\theta}) = 0 \implies$ unbiased estimator
2. Efficiency (有效性):
   Under same sample size, if $D(\hat{\theta}_1) < D(\hat{\theta}_2) \Rightarrow \hat{\theta}_1$ is more efficient than $\hat{\theta}_2$.
3. Consistency (相合性，一致性):
   Def. If $\lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1 \Rightarrow$ consistent estimator
   where ε is an arbitrarily small positive number.

Note: if $\text{Bias}(\hat{\theta}) = 0$ & $D(\theta) \to 0$ as $n \to \infty$, then $\hat{\theta}$ is a consistent estimator of θ.

Quick check of estimator

Biasedness (无偏性): $E(\hat{\theta}) = \theta$

Consider drawing a sample of n i.i.d. R.V.s. $X_1, X_2, \dots, X_n$ from population X with μ and σ^2, show that

1. $\hat{\mu}_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of μ.
2. $\hat{\mu}_2 = \sum_{i=1}^{n} c_i X_i$ is an unbiased estimator of μ, where $\sum_{i=1}^{n} c_i = 1, c_i > 0, i = 1,2,\cdots,n$.
3. $\hat{\sigma}_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an unbiased estimator of σ^2.
4. $\hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is not an unbiased estimator of σ^2.

Quick Check

Ex. Let $X_1, X_2, \dots, X_{16}$ be a sample from $X \sim \mathcal{N}(\mu, \sigma^2)$. If $c\sum_{i=1}^{15}(X_{i+1} - X_i)^2$ is an unbiased estimator of σ^2, then $c =$____________.

Sol.

Let $X = (X_{i+1} - X_i) \sim \mathcal{N}(0, 2\sigma^2), i = 1,2,\dots,15$

If $c\sum_{i=1}^{15}(X_{i+1} - X_i)^2$ is an unbiased estimator, it satisfies

$$E\left[c\sum_{i=1}^{15}(X_{i+1} - X_i)^2\right] = E\left(c\sum_{i=1}^{15} X^2\right) = c\sum_{i=1}^{15} E(X^2)$$
$$= 15c \cdot E(X^2) = 15c[D(X) + E(X)^2] = 15c[2\sigma^2 + 0] = \sigma^2$$

Therefore, $c = \frac{1}{30}$.

Quick check of estimator

Efficiency (有效性): $D(\hat{\theta}_1) < D(\hat{\theta}_2) \Rightarrow \hat{\theta}_1$ is more efficient

Consider drawing a sample of n i.i.d. R.V.s. $X_1, X_2, \ldots, X_n$ from population X with μ and σ^2, show that $\hat{\mu}_1 = \frac{1}{2}X_1 + \frac{1}{2}X_2$ is more efficient than $\hat{\mu}_2 = \frac{1}{4}X_1 + \frac{3}{4}X_2$.

Proof:

$$D(\hat{\mu}_1) = D\left(\frac{1}{2}X_1 + \frac{1}{2}X_2\right) = \frac{1}{4}D(X_1) + \frac{1}{4}D(X_2) = \frac{1}{2}\sigma^2$$

$$D(\hat{\mu}_2) = D\left(\frac{1}{4}X_1 + \frac{3}{4}X_2\right) = \frac{1}{16}D(X_1) + \frac{9}{16}D(X_2) = \frac{5}{8}\sigma^2$$

$D(\hat{\mu}_1) < D(\hat{\mu}_2) \quad \Rightarrow \quad \hat{\mu}_1$ is more efficient than $\hat{\mu}_2$

Generalizable to $\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\hat{\mu}_2 = \sum_{i=1}^{n} c_i X_i$.

Quick check of estimator

Consistency (相合性，一致性):

$$\lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1$$

Consider drawing a sample of n i.i.d. R.V.s. $X_1, X_2, \dots, X_n$ from population X with μ and σ^2, show that

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

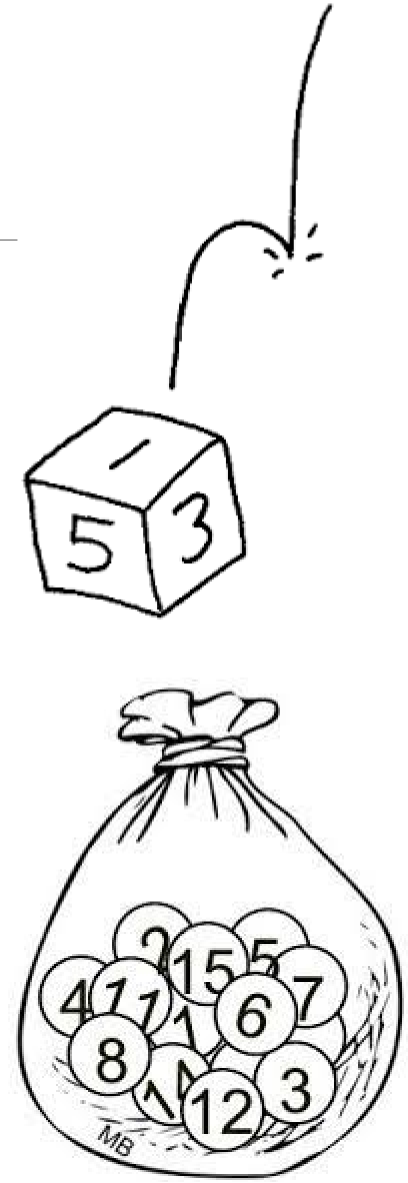
is a consistent estimator of μ.

Proof: From law of large number, $\lim_{n\to\infty} P\{|\hat{\mu} - \mu| < \varepsilon\} = 1$,

Therefore, $\hat{\mu}$ is a consistent estimator of μ.

Lecture 17

- Moment Estimator
- Estimator Selection Criteria
- Confidence Intervals



Data → Parameters

Parameter estimation

Point estimation
点估计

Interval estimation
区间估计

Provide a specific value to estimate θ.

But missing important information:

- Accuracy
- The error range
- The credibility

Fishing in a murky lake with a spear

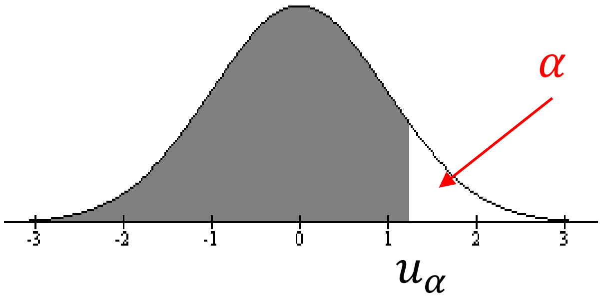
Find an interval (A, B) to approach θ.

The length of $(A, B) \Rightarrow$ how close we can estimate θ.

Fishing with a net

Quick Check

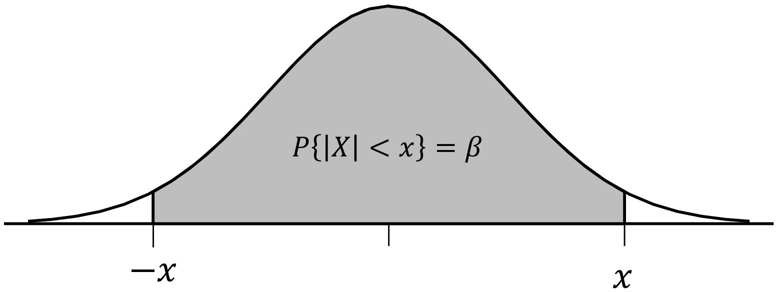
Ex. For R.V. $X \sim \mathcal{N}(0,1)$, we define u_α as $P\{X > u_\alpha\} = \alpha$, where $0 \le \alpha \le 1$.



If $P\{|X| < x\} = \beta$, then the value of x should be

(A) $u_{\frac{\beta}{2}}$ (B) $u_{1-\frac{\beta}{2}}$ (C) $u_{\frac{1-\beta}{2}}$ (D) $u_{1-\beta}$

C



Preliminary on Probability Intervals

For a PDF of $X \sim \mathcal{N}(0,1)$, 95% of its area is within $(-1.96, 1.96)$.

For a PDF of $X \sim \mathcal{N}(0, \sigma^2)$, 95% of its area is within $(-1.96 \cdot \sigma, 1.96 \cdot \sigma)$.

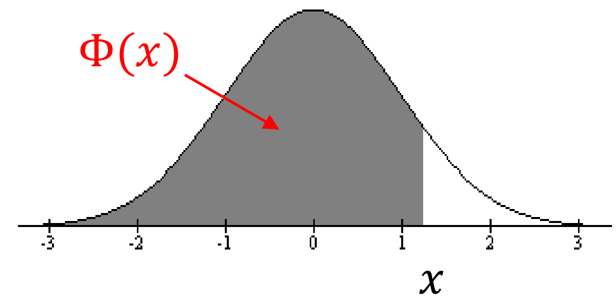
According to CLT., $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$, 95% of its area is within $\left(\mu - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, \mu + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right)$.

$\Rightarrow$ For 95% probability, $\bar{X} \in \left(\mu - 1.96 \frac{\sigma}{\sqrt{n}}, \mu + 1.96 \frac{\sigma}{\sqrt{n}}\right)$.

Φ has been numerically computed

Table 5.1 Area $\Phi(x)$ Under the Standard Normal Curve to the Left of X.

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

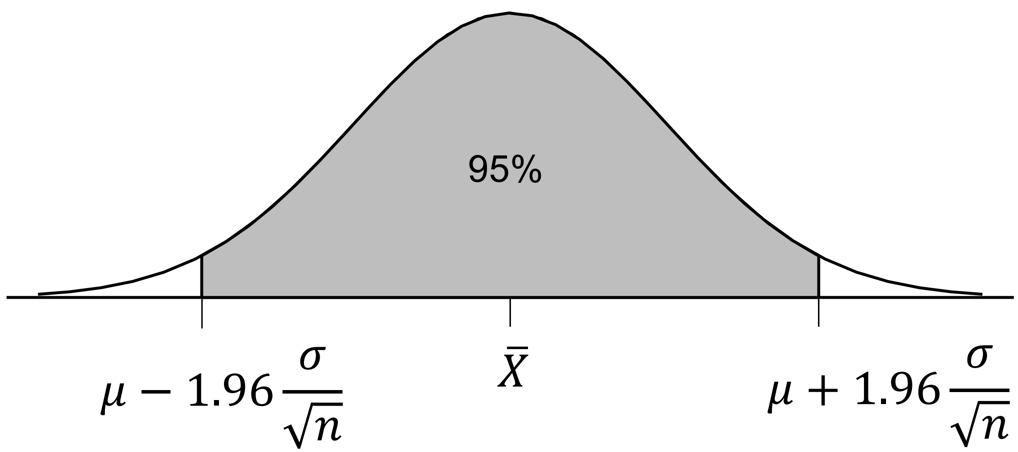


Given x, obtain $\Phi(x)$ numerically from the table.

Example on confidence interval (置信区间)

From CLT., $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ as $n \to \infty$. For a normal distribution, 95% of its area is within $\pm 1.96 \frac{\sigma}{\sqrt{n}}$ from the center.

For 95% probability, $\bar{X} \in \left(\mu - 1.96 \frac{\sigma}{\sqrt{n}}, \mu + 1.96 \frac{\sigma}{\sqrt{n}}\right)$.



Example on confidence interval

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

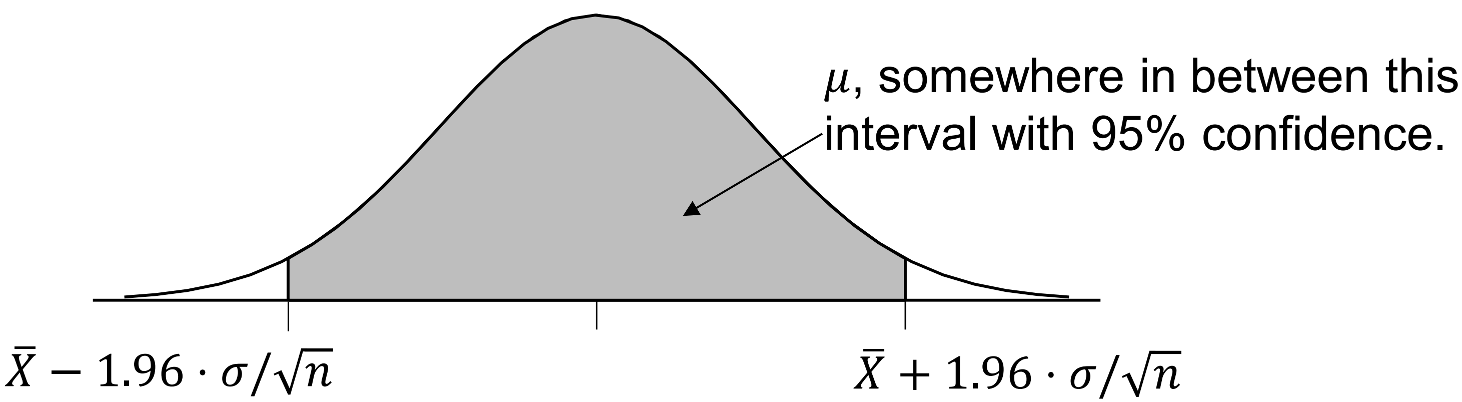
For 95% probability, $\bar{X} \in \left(\mu - 1.96\frac{\sigma}{\sqrt{n}}, \mu + 1.96\frac{\sigma}{\sqrt{n}}\right)$.

$$\mu - 1.96\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1.96\frac{\sigma}{\sqrt{n}}$$

That is: For 95% probability,

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

We call $\boldsymbol{\bar{X} \pm 1.96 \cdot \sigma/\sqrt{n}}$ **a 95% confidence interval** for μ.



Confidence intervals

Consider drawing a sample of n i.i.d. R.V.s. $X_1, X_2, \dots, X_n$ from population $X \sim F(x;\theta)$, where θ is unknown, we have an interval $\underline{\theta} < \theta < \overline{\theta}$

$$P\{\underline{\theta} < \theta < \overline{\theta}\} \geq 1 - \alpha$$

<u>Def.</u>

$(\underline{\theta}, \overline{\theta})$: a confidence interval for θ with confidence level $1 - \alpha$.

(置信区间) (置信水平)

where $\underline{\theta}, \overline{\theta}$ are the lower and upper bound of the interval.

We want:

1. A high accuracy $\Rightarrow$ a small $\overline{\theta} - \underline{\theta}$;
2. A high confidence level $\Rightarrow$ a large $P\{\underline{\theta} < \theta < \overline{\theta}\}$.

Two contradictory goals!

Ex. A travel agency surveys the average spending of local tourists. It randomly samples 100 tourists. The average spending is $\bar{x} = 80$ dollars. According to our experience, the spending of local tourists follow normal distribution with standard deviation $\sigma = 12$. Find the 95% confidence interval of parameter μ.

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

Sol.

The confidence interval of μ is given by

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$

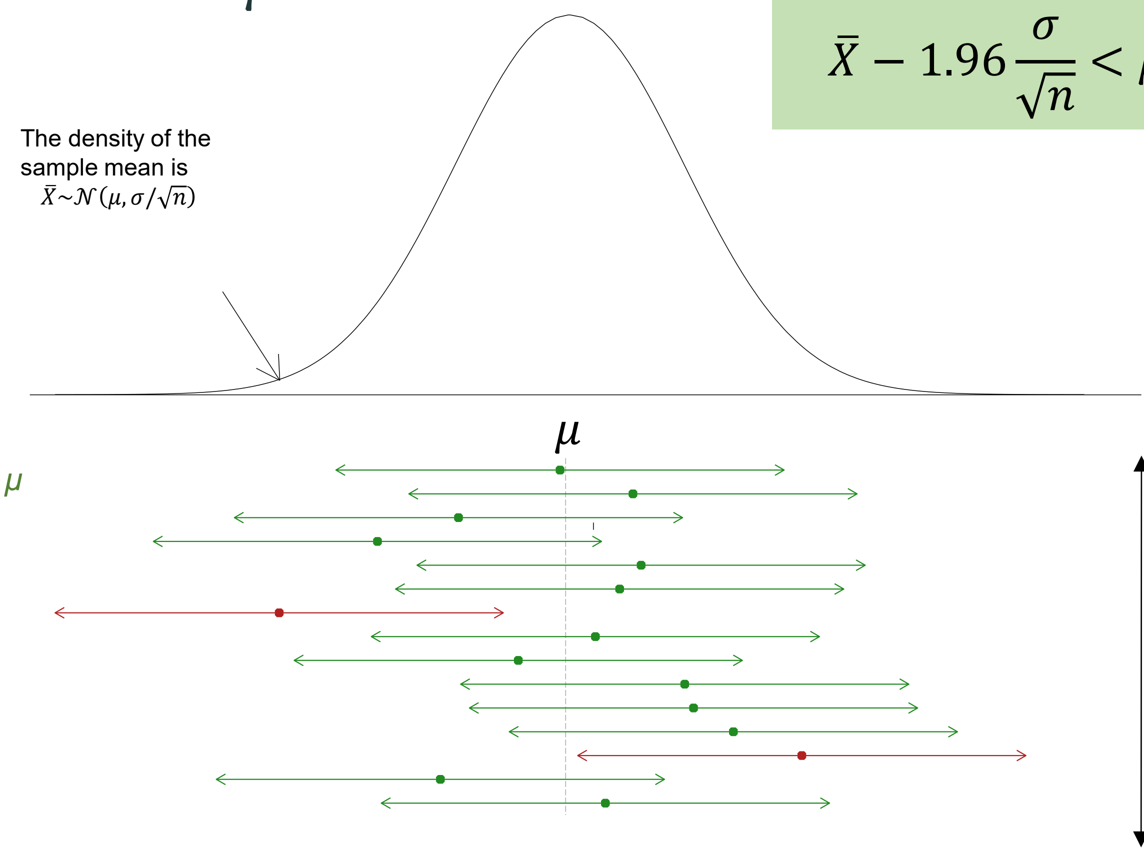
where $\bar{x} = 80$, $\sigma = 12$ and $n = 100$

$$77.6 < \mu < 82.4$$

What does "95% confidence" really mean?

About 95% of the constructed intervals cover the true population mean μ.

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$



green CIs cover μ

red CIs miss μ

95% of the intervals covering μ

True or False

95% confidence interval of average spending: $77.6 < \mu < 82.4$

True or False and explain: We are 95% confident that the average spending in this sample have been in between 77.6 and 82.4.

False. *The confidence interval* $\bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}}$ *definitely* (100%) *contains the sample mean* $\bar{X}$*, not just with probability* 95%.

True or False and explain: 95% of the spending of tourists have been in 77.6 to 82.4.

False. *The confidence interval is for covering the population mean* μ*, not for covering 95% of the entire population.*

True or False

95% confidence interval of average spending: $77.6 < \mu < 82.4$

True or False and explain: If a new random sample of size 100 is taken, we are 95% confident that the new sample mean will be between 77.6 and 82.4.

False. The confidence interval is for covering the population mean μ, not for covering the mean of another sample.