UENF
Universidade Estadual do Norte Fluminense Darcy Ribeiro

CCT
LCMAT

Inferencia Estatística: Distribuições Amostrais e Intervalos de Confiança

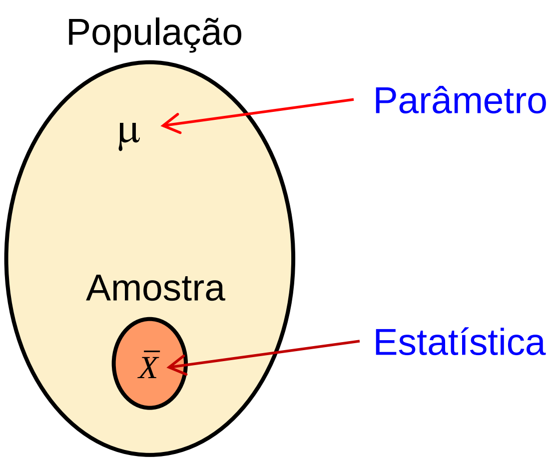
Prof. Fermín Alfredo Tang Montané

Inferência Estatística
Definições

O objetivo principal da inferência estatística é obter informações sobre determinada característica da população baseando-se apenas das informações obtidas de uma amostra.

Parâmetro: quantidades da população, em geral, desconhecidas e sobre as quais temos interesse. Representados por letras gregas: μ, σ, etc.

Estatística: quantidades calculadas com base nos elementos da amostra. Representadas por letras do alfabeto latino: $\bar{X}$, s, etc.



Estimador: é uma estatística destinada a estimar um parâmetro de interesse da população. Por exemplo: $\bar{X}$ é um estimador de μ.

Estimativa: é o valor numérico do estimador.

Denominação	Estimador	Parâmetro
Média	$\bar{X}$	μ
Variância	S^2	σ^2
Número de elementos	n	N
Proporção	$\hat{p}$	p

Inferência Estatística

Definições

Vício ou viesado: um estimador é não viciado ou não viesado para um parâmetro se $E(\hat{\theta})=\theta$,ou seja, se a esperança matemática é igual ao valor do parâmetro.

Consistência: um estimador é consistente, se à medida que o tamanho da amostra aumenta, sua esperança matemática converge para o parâmetro de interesse e sua variância converge para zero.

Eficiência: dados dois estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$, não viciados para um parâmetro, dizemos que $\hat{\theta}_1$ é mais eficiente que $\hat{\theta}_2$ se $V(\hat{\theta}_1)<V(\hat{\theta}_2)$.

Distribuição Amostral
Definições

- As estatísticas e os parâmetros são funções de variáveis aleatórias.
- As estatísticas e os parâmetros são também variáveis aleatórias, por tanto, ambas possuem distribuição de probabilidade, esperança matemática e variância.
- Distribuições amostrais são distribuições de probabilidade para estatísticas amostrais que servem como estimadores de parâmetros populacionais. p. e. distribuição da média amostral $\bar{X}$.

- Estudam-se as distribuições amostrais para:
  - Média;
  - Variância;
  - Proporção;
  - Diferência de duas médias
  - Diferencia de duas proporções

Distribuição Amostral da Média

Exemplo

Considere uma população em que a variável aleatória X assume os valores do conjunto {1,3,5,5,7}. A distribuição de probabilidade de X é dada por:

X = x	**1**	**3**	**5**	**7**
P(X = x)	1/5	1/5	2/5	1/5

$$\mu = E(X) = 1.1/5 + 3.1/5 + 5.2/5 + 7.1/5 = 4{,}2$$

$$\sigma^2 = V(X) = (1-4{,}2)^2.1/5 + (3-4{,}2)^2.1/5 + \dots + (7-4{,}2)^2.1/5 = 4{,}16$$

Considere todas as amostras possíveis de tamanho n=2, selecionadas ao acaso e com reposição dessa população, e encontre a distribuição da média amostral $\bar{X}$, onde:

$$\bar{X} = \frac{X_1 + X_2}{2}$$

onde:

X_1: valor selecionado na primeira extração;

X_2: valor selecionado na segunda extração.

Distribuição Amostral da Média

Exemplo

Amostra (X_1,X_2)	Probabilidade	Média Amostral
(1,1)	1/25	1
(1,3)	1/25	2
(1,5)	2/25	3
(1,7)	1/25	4
(3,1)	1/25	2
(3,3)	1/25	3
(3,5)	2/25	4
(3,7)	1/25	5
(5,1)	2/25	3
(5,3)	2/25	4
(5,5)	4/25	5
(5,7)	2/25	6
(7,1)	1/25	4
(7,3)	1/25	5
(7,5)	2/25	6
(7,7)	1/25	7
	1	

Distribuição Amostral da Média
Exemplo

A distribuição de probabilidade para a média amostral $\bar{X}$ é:

$\bar{x}$	1	2	3	4	5	6	7
$P(\bar{X}=\bar{x})$	1/25	2/25	5/25	6/25	6/25	4/25	1/25

$$\mu = E(\bar{x}) = 1.(1/25) + 2.(2/25) + \ldots + 7.(1/25) = 4,2$$

$$\sigma^2 = V(\bar{x}) = (1 - 4,2)^2.1/25 + \ldots + (7 - 4,2)^2.1/25 = 2,08$$

$$\sigma^2 = V(\bar{x}) = 4,16/2$$

Distribuição Amostral da Média

Exemplo

Repetindo o mesmo procedimento para amostras de tamanho n = 3, tem-se a seguinte distribuição de probabilidade para a média amostral $\bar{X}$:

$\bar{x}$	$P(\bar{X} = \bar{x})$
1	1/125
5/3	3/125
7/3	9/125
3	16/125
11/3	24/125
13/3	27/125
5	23/125
17/3	15/125
19/3	6/125
7	1/125

$E(\bar{x}) = 1.(1/125) + \ldots + 7.(1/125)$

$E(\bar{x}) = 4{,}2$

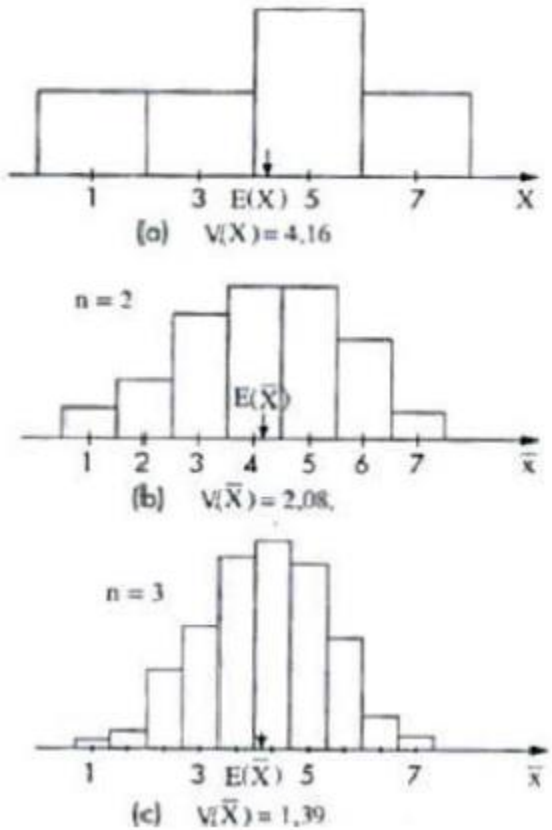
$V(\bar{x}) = (1-4{,}2)^2.1/125 + \ldots + (7-4{,}2)^2.1/125$

$V(\bar{x}) = 1{,}39 = 4{,}16/3$

Distribuição Amostral da Média

Ilustração

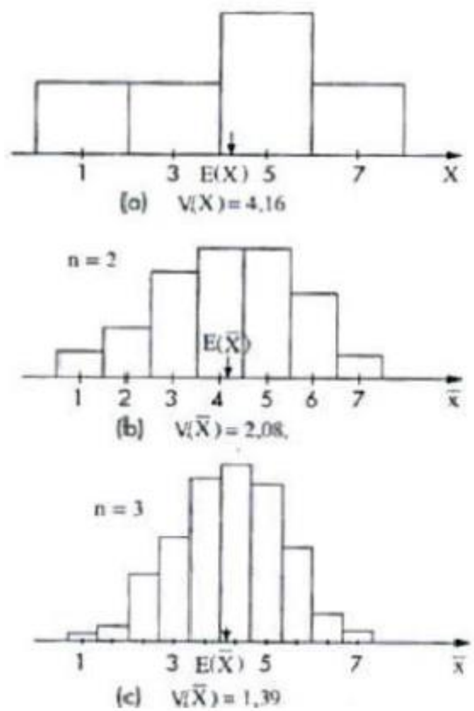
Os histogramas correspondentes da variável aleatória X e da variável aleatória para n = 2 e n = 3 estão apresentados abaixo:



Distribuição Amostral da Média
Propriedades

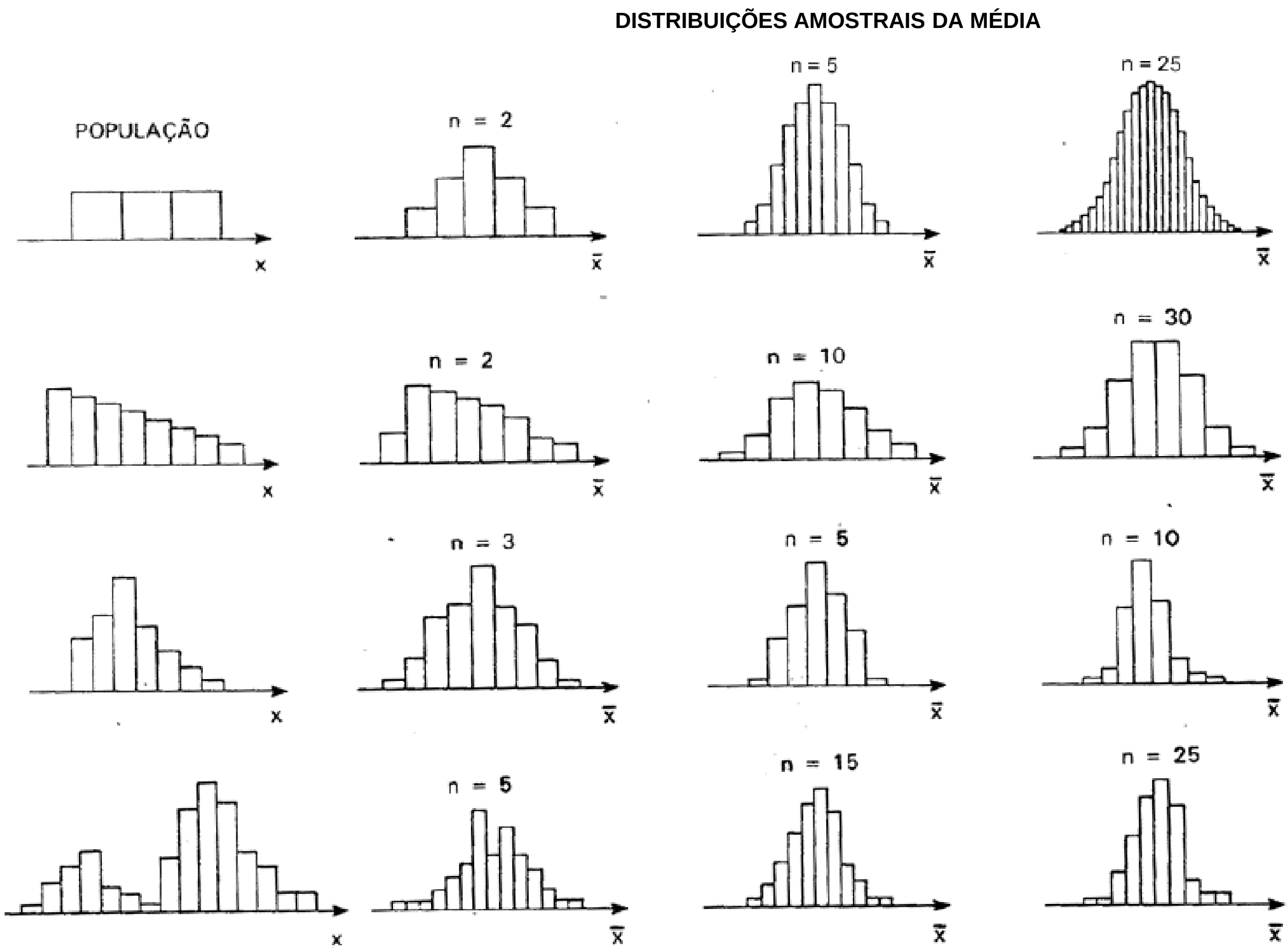
Dos histogramas observa-se que:

- Conforme n aumenta os valores da média amostral tendem a se concentrar cada vez mais em torno da E(X), pois a variância diminui;
- Os valores extremos passam a ter pequenas probabilidades de ocorrência;
- Conforme n aumenta, a forma da distribuição das médias se aproxima da distribuição normal;



Distribuição Amostral da Média

Ilustração



Fonte: BUSSAB & MORETTIN. ***Estatística Básica***. São Paulo, Atual, 3ª edição, 1986, pp. 197.

Distribuição Amostral da Média

Exemplo

Teorema do Limite Central

Seja X uma v. a. que tem média μ e variância σ^2. Para amostras $X_1, X_2, \ldots, X_n$, retiradas ao acaso e com reposição de X, a distribuição de probabilidade da média amostral $\overline{X}$ *aproxima-se*, *para n grande*, de uma distribuição normal, com média μ e variância σ^2 / n, ou seja,

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \text{ para } n \text{ grande}$$

O desvio padrão $\sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ é denominado ***erro padrão da média.***

Distribuição Amostral da Média

Definição

Considerando-se que X tem média μ e variancia σ^2, segue-se que, a distribuição da média amostral $\bar{X}$ é dada por:

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

Notação:

Para distinguir entre as médias e varianças populacionais e amostrais as vezes é conveniente colocar a variável populacional ou amostral como sub-indice.

μ_X e σ_X^2 : Média e variância populacional (ou simplesmente μ e σ^2).

$\mu_{\bar{X}}$ e $\sigma_{\bar{X}}^2$: Média e variância amostral.

Assim, a expressão acima pode ser escrita como: $\bar{X} \sim N\left(\mu_{\bar{X}}, \sigma_{\bar{X}}^2\right)$

Distribuição Amostral da Média

Exemplos

Exemplo1:

Uma variável aleatória X assume os valores 3, 6 e 8 com, respectivamente, probabilidades 0,4; 0,3 e 0,3. Uma amostra de 40 observações com reposição é obtida aleatoriamente. Qual a probabilidade da média amostral ser maior que 5?

Resposta:

Como o tamanho da amostra n=40 é suficientemente grande é possível aplicar o teorema do limite central. Assim, a media amostral tem distribuição normal:

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

onde: $\mu = \mu_X = E(X)$ e $\sigma^2 = \sigma_X^2 = V(X)$

Calcula-se, por definição: $E(X) = (3)(0,4) + (6)(0,3) + (8)(0,3) = 5,4$

$$V(X) = (3-5,4)^2(0,4) + (6-5,4)^2(0,3) + (8-5,4)^2(0,3) = 4,44$$

Pede-se calcular: $P(\bar{X} > 5)$ Sabe-se que: $\bar{X} \sim N\left(5,4; \frac{4,44}{40}\right) = N(5,4;\ 0,111)$

Distribuição Amostral da Média
Exemplos

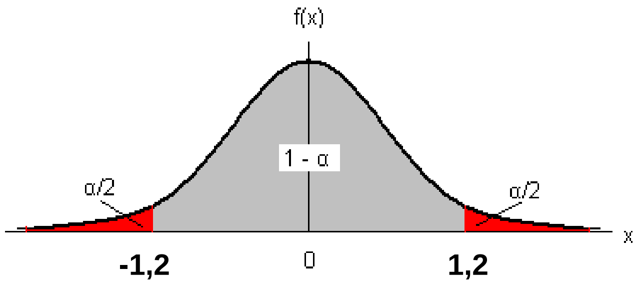
Resposta (Continuação):

Pede-se calcular: $P(\bar{X} > 5)$ Sabe-se que: $\bar{X} \sim N\left(5,4; \frac{4,44}{40}\right) = N\left(5,4;\ 0,111\right)$

Padronizar a variável de maneira a ter distribuição normal padrão N(0,1):

$$P(\bar{X} > 5) = P\left(\frac{\bar{X}-\mu}{\sigma} > \frac{5-\mu}{\sigma}\right) = P\left(z > \frac{5-5,4}{0,333}\right) = P\left(z > -1,2\right)$$

Utiliza-se a distribuição normal padrão de maneira a calcular a probabilidade correspondente a o valor -1,2. Considerando-se a simetria da normal, temos que:



$$P(\bar{X} > 5) = P\left(z > -1,2\right) = 0,5 + P(0 < z < 1,2)$$

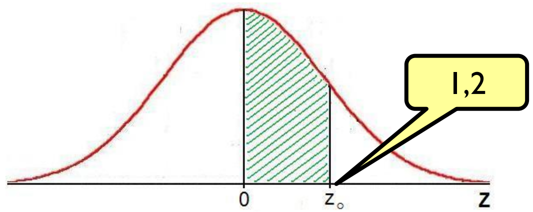
Utilizando a tabela normal se obtêm:

$$\boxed{P(\bar{X} > 5) = 0,5 + 0,3849 = 0,8849}$$

Distribuição Amostral da Média

Exemplos

Resposta: Da tabela normal.



$$P(0 < z < z_o) = \int_0^{z_o} \frac{1}{\sigma.\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2} dz$$

Define-se:

$Z_o = 1{,}2$

Procura-se:

$P(0 < z < 1{,}2) = 0{,}3849$

z_0	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,49865	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49897	0,49900
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,49929
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,49950
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,49965
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,49976
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,49983
3,6	0,49984	0,49985	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,49989
3,7	0,49989	0,49990	0,49990	0,49990	0,49991	0,49991	0,49992	0,49992	0,49992	0,49992
3,8	0,49993	0,49993	0,49993	0,49994	0,49994	0,49994	0,49994	0,49995	0,49995	0,49995
3,9	0,49995	0,49995	0,49996	0,49996	0,49996	0,49996	0,49996	0,49996	0,49997	0,49997

Distribuição Amostral da Média

Exemplos

Exemplo2:

O faturamento diário de um supermercado está normalmente distribuído com média de R$20.000,00 e desvio-padrão de R$2000,00. Qual a probabilidade do faturamento ultrapassar R$1230000,00 em 60dias?

Resposta:

O faturamento diário deve ser considerado como a variável populacional X, com parâmetros μ e σ conhecidos. $\mu = 20.000$ e $\sigma = 2.000$

Uma amostra de tamanho n=60 é considerada, com faturamentos para 60 dias. O faturamento médio nos 60 dias corresponde a média amostral. Sabe-se que tem distribuição normal:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N(20.000; 66.666,66)$$

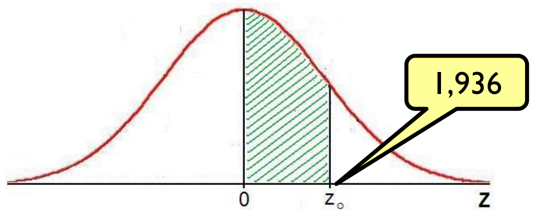
Pede-se: $P(\bar{X} > \frac{1.230.000}{60}) = P(\bar{X} > 20.500)$ Padronizar a variável como normal N(0,1):

$$P(\bar{X} > 20.500) = P\left(z > \frac{20.500 - 20.000}{258,198}\right) = P(z > 1,936)$$

Distribuição Amostral da Média

Exemplos

Resposta: Da tabela normal.



Define-se: $Z_o = 1{,}936$

Procura-se: $P(0 < z < 1{,}936) = 0{,}4738$

Finalmente:

$$P(\bar{X} > 20.500) = P\left(z > 1{,}936\right)$$
$$= 0{,}5 - 0{,}4738 = 0{,}0262$$

z_o	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,49865	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49897	0,49900
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,49929
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,49950
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,49965
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,49976
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,49983
3,6	0,49984	0,49985	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,49989
3,7	0,49989	0,49990	0,49990	0,49990	0,49991	0,49991	0,49992	0,49992	0,49992	0,49992
3,8	0,49993	0,49993	0,49993	0,49994	0,49994	0,49994	0,49994	0,49995	0,49995	0,49995
3,9	0,49995	0,49995	0,49996	0,49996	0,49996	0,49996	0,49996	0,49996	0,49997	0,49997

Distribuição da Média Amostral

Exemplos

Exemplo3:

Considere que a distribuição dos níveis de colesterol para todos os homens de 20 a 74 anos está normalmente distribuído com média 211 mg e desvio-padrão de 46mg. Selecionando 25 homens desta população, determine:

a) A proporção destes 25 homens que terá um valor médio inferior a 230mg;

b) O valor médio de nível de colesterol que limita os10% dos valores mais baixos da distribuição amostral;

c) Os limites superior e inferior que incluem 95% das médias das amostras de tamanho25;

d) Qual deve ser o tamanho das amostras para que 95% de suas médias se encontrem a ±5mg da média da população?

Distribuição Amostral da Média

Caso com Variância Conhecida e População Finita

Se uma amostra de tamanho n é retirada (sem reposição) de uma **população finita** de tamanho N (N>n), utiliza-se o **fator de correção** para a variância.

$$\boxed{\frac{N-n}{N-1}}$$

De maneira que: $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\frac{N-n}{N-1}$

Como: $\bar{X} \sim N\left(\mu_{\bar{X}}, \sigma_{\bar{X}}^2\right)$

Então: $\boxed{\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\frac{N-n}{N-1}\right)}$

Distribuição Amostral da Média

Exemplo

Exemplo:

As lâmpadas fabricadas por uma indústria tem duração média de 800 horas e desvio-padrão de 100 horas. É escolhida aleatoriamente 200 lâmpadas de um lote de 2000 lâmpadas. Determine a probabilidade da média destas lâmpadas escolhidas ser superior a 810 horas.

Distribuição Amostral de uma Proporção

Definição

Considere que p é a proporção de elementos de uma população que possui uma certa característica. Para cada elemento da população pode ser definida uma variável aleatória X, tal que:

$$X = \begin{cases} 1, & \text{se o indivíduo apresenta a caraterística} \\ 0, & \text{caso contrário.} \end{cases}$$

Assim, X é uma variável aleatória com distribuição Bernoulli. Sabe-se que:

$$E(X) = p \text{ e } V(X) = pq$$

Considere agora que uma amostra x_1, x_2, ..., x_n é extraída da população de maneira aleatória como reposição. Cada varíavel do tipo descrito anteriormente. Seja:

$$S_n = X_1 + X_2 + \ldots + X_n$$

Onde S_n, representa o número total de individuos que apresentam a característica na amostra. Observe que S_n, tem distribuição binomial, onde:

$$E(S_n) = np \text{ e } V(S_n) = npq$$

Distribuição Amostral de uma Proporção

Definição

A proporção amostral é definida como:

$$\boxed{\hat{p} = \frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}}$$

Como S_n tem distribuição binomial e n é uma constante, aplican-se as propriedades do valor esperado e a variância.

$$E(\hat{p}) = E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{np}{n} = p$$

$$V(\hat{p}) = V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}V(S_n) = \frac{npq}{n^2} = \frac{pq}{n}$$

Observe que $\hat{p}$ é um estimador não viciado e consistente para p, uma vez que:

$$E(\hat{p}) = p$$

$$\lim_{n\to\infty} V(\hat{p}) = \lim_{n\to\infty} \frac{pq}{n} = 0$$

Distribuição Amostral de uma Proporção

Definição

Pelo teorema do limite central sabe-se que $\hat{p}$ tem distribuição normal, para n suficientemente grande:

$$\boxed{\hat{p} \sim N\left(p, \frac{pq}{n}\right)}$$

Padronizando a variável de maneira a ter distribuição normal padrão:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$ onde: $z \sim N(0,1)$

Lembre que: $q = (1 - p)$

Distribuição Amostral de uma Proporção

Exemplos

1) A proporção de peças defeituosas de um lote é de 40%. Foi coletada aleatoriamente uma amostra de 30 peças com reposição. Qual a probabilidade desta amostra fornecer uma proporção de peças defeituosas menor que 50%?

2) Qual a probabilidade de ocorrer entre 40% e 50% de caras em 120 lançamentos de uma moeda não viciada?

Distribuição Amostral da Diferença de Duas Médias

Sejam duas populações 1 e 2, com médias μ_1 e μ_2 e desvios-padrão σ_1 e σ_2, respectivamente. São retiradas, independentemente, amostras de tamanho n_1 da população 1 e de tamanho n_2 da população 2. De todas as possíveis amostras retiradas pode-se obter a distribuição amostral da diferença entre as duas médias. Se n_1 e n_2 forem suficientemente grandes:

A distribuição amostral da diferença das médias é dada por:

$$\boxed{(\bar{x}_1 - \bar{x}_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}$$

Padronizando a variável de maneira a ter distribuição normal padrão:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

onde: $z \sim N(0,1)$

Distribuição Amostral da Diferença de Duas Médias

Exemplo:

As lâmpadas elétricas do fabricante A têm duração média de 1400horas, com desvio-padrão de 200 horas, enquanto as do fabricante B têm duração média de 1200 horas, com desvio-padrão de 100 horas. Se forem ensaiadas amostras aleatórias de 125 lâmpadas de cada marca, qual será a probabilidade de que as lâmpadas da marca A tenham vida média maior do que as da marca B de pelo menos 160 horas?

Distribuição Amostral da Diferença de Proporções

Considere que seja extraídas amostras de tamanho n_1 da população1cuja proporção de elementos com uma determinada característica seja p_1 e que sejam extraídas amostras de tamanho n_2 da população 2 cuja proporção de elementos com a referida característica seja p_2. A distribuição amostral da diferença das duas proporções é dada por:

A distribuição amostral da diferença das duas proporções é dada por:

$$\boxed{(\hat{p}_1 - \hat{p}_2) \sim N\left(p_1 - p_2, \frac{p_1.q_1}{n_1} + \frac{p_2.q_2}{n_2} \right)}$$

Padronizando a variável de maneira a ter distribuição normal padrão:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1.q_1}{n_1} + \frac{p_2.q_2}{n_2}}}$$

onde: $z \sim N(0,1)$

Distribuição Amostral da Diferença de Proporções

Exemplo:

Duas pessoas A e B jogam uma partida do tipo "cara e coroa" onde cada uma lança 50 vezes uma moeda não viciada. O jogador A vencerá o jogo se conseguir 5 ou mais caras do que o jogador B e, se isso não ocorrer, o jogador B vencerá. Determine a probabilidade de cada jogador ganhar.

Intervalos de Confiança

Podem ser determinados intervalos de confiança para diferentes parâmetros populacionais como:

- Média
- Variância
- Proporção
- Diferença de duas médias
- Diferença de duas proporções

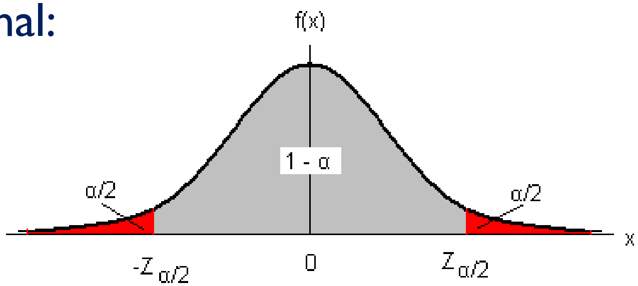
O intervalo de confiança para uma parâmetro populacional é baseado na distribuição amostral de seu estimador amostral assim como em um nível de confiança desejado.

O intervalo de confiança para μ é baseada na distribuição da média amostral e no nível de confiança.

Intervalo de Confiança da Média
com Variância Conhecida

No caso com variância conhecida população infinita. Pelo Teorema do Limite Central, sabe-se que a media amostral tem distribuição Normal:

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



Normalizando, se tem que:

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Como a variável tem distribuição normal, define-se os valores críticos $-Z_{\alpha/2}$ e $Z_{\alpha/2}$ para um nível de confiança ($1-\alpha$):

$$-Z_{\alpha/2} \le \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \le Z_{\alpha/2}$$

Escrevendo as desigualdades em função de μ:

$$-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{x}-\mu \qquad \bar{x}-\mu \le Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Assim, o intervalo de confiança para μ é:

$$\boxed{\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$$

Sendo que o nível de confiança ($1-\alpha$), corresponde a probabilidade de que a variância σ^2 se encontre no intervalo calculado.

$$\boxed{P\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha}$$

Intervalo de Confiança da Media
com Variância Conhecida

Exemplo:

Uma variável aleatória contínua X tem média μ e variância $\sigma^2=16$. Uma Amostra é obtida aleatoriamente com 40 valores obtendo média igual a13. Construir um intervalo de confiança com nível de significância de 5% para μ.

Resposta:

O intervalo de confiança para a média é dado por:

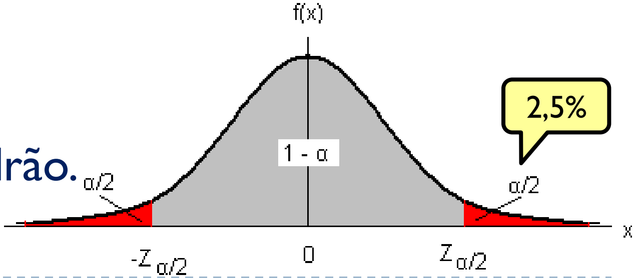
$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Sabe-e que: $n = 40;\ \bar{x} = 13;\ \sigma = 4.$

Calcula-se os valores críticos $-Z_{\alpha/2}$ e $Z_{\alpha/2}$ para a distribuição normal padrão com $(1-\alpha)=0,95$ e $\alpha=0,05$. Dada a simetria da normal somente é preciso calcular um valor crítico, sendo que:

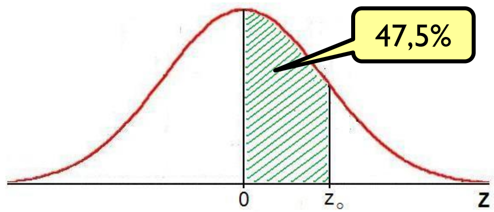
$$Z_{\alpha/2} = Z_{0,05/2} = Z_{0,025}$$

Os valores críticos são obtidos da tabela normal padrão.



Intervalo de Confiança da Media
com Variância Conhecida

Resposta: Da tabela normal.



$$P(0 < z < z_0) = \int_0^{z_0} \frac{1}{\sigma.\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

z_0	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,49865	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49897	0,49900
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,49929
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,49950
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,49965
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,49976
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,49983
3,6	0,49984	0,49985	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,49989
3,7	0,49989	0,49990	0,49990	0,49990	0,49991	0,49991	0,49992	0,49992	0,49992	0,49992
3,8	0,49993	0,49993	0,49993	0,49994	0,49994	0,49994	0,49994	0,49995	0,49995	0,49995
3,9	0,49995	0,49995	0,49996	0,49996	0,49996	0,49996	0,49996	0,49996	0,49997	0,49997

$$Z_{\alpha/2} = Z_{0,475} = 1,96$$

Considerando: $n = 40;\ \bar{x} = 13;\ \sigma = 4.$

O intervalo de confiança para a média é:

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$13 - (1,96)\frac{40}{\sqrt{40}} \le \mu \le 13 + (1,96)\frac{40}{\sqrt{40}}$$

$$\boxed{0,60387 \le \mu \le 25,3961}$$

Intervalo de Confiança para a Média
com Variância Conhecida e População Finita

Intervalos de Confiança

Para população finita utiliza-se o **fator de correção** para o desvio-padrão:

$$\boxed{\sqrt{\frac{N-n}{N-1}}}$$

Assim, o intervalo de confiança para μ é:

$$\boxed{\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$

Sendo que o nível de confiança (1-α), corresponde a probabilidade de que a média μ se encontre no intervalo calculado.

$$\boxed{P\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right) = 1-\alpha}$$

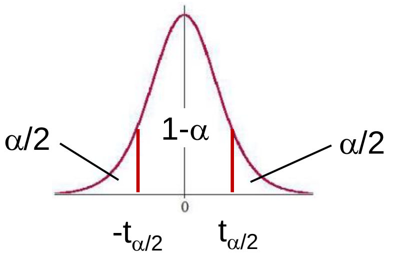
Intervalo de Confiança para a Média
com Variância Desconhecida

A variância populacional é estimada através da variância amostral:

$$\boxed{s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

O intervalo de confiança é obtido através da distribuição t de Student com (n–1) graus de liberdade.

$$-t_{\alpha/2} \le \frac{\overline{x}-\mu}{s/\sqrt{n}} \le t_{\alpha/2}$$



Escrevendo as desigualdades em função de μ: $-t_{\alpha/2}\frac{s}{\sqrt{n}} \le \overline{x}-\mu$ $\qquad$ $\overline{x}-\mu \le t_{\alpha/2}\frac{s}{\sqrt{n}}$

Assim, o intervalo de confiança para μ é:

$$\boxed{\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}}$$

Sendo que o nível de confiança (1-α), corresponde a probabilidade de que a média μ se encontre no intervalo calculado.

$$\boxed{P\left(\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) = 1-\alpha}$$

Intervalo de Confiança para a Média
com Variância Desconhecida

Exemplo:

Uma amostra de 14 alunos de uma turma de uma determinada escola tiveram média 6,8 e variância 4 numa prova. Construir um intervalo de confiança de 90% para a média das notas dos alunos desta escola.

Resposta:

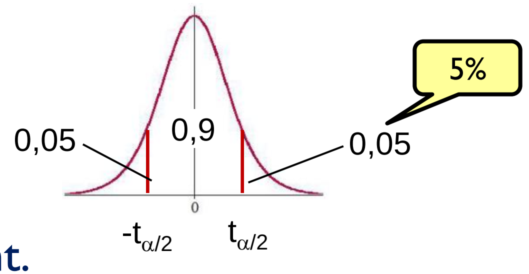
O intervalo de confiança para a média é dado por:

$$\boxed{\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}}$$

Sabe-e que: $n = 14;\ \bar{x} = 6{,}8;\ \mathrm{s}^2 = 4.$

Calcula-se os valores críticos $-t_{\alpha/2}$ e $t_{\alpha/2}$ para a distribuição t-student com (n-1)=13 graus de liberdade e $(1-\alpha)=0{,}90$ e $\alpha=0{,}10$. Dada a simetria da t-sudent somente é preciso calcular um valor crítico, sendo que:

$$t_{\alpha/2} = \mathrm{t}_{0,10/2} = t_{0,05}$$

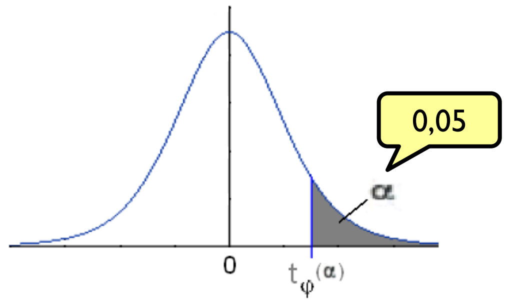


Os valores críticos são obtidos da tabela t-student.

Intervalo de Confiança da Media

Variância Desconhecida

Resposta: Da tabela t-student.



$t_{\varphi}(\alpha) = \quad t_{13}(0,05) = 1,771$

Considerando: $n = 14;\ \bar{x} = 6,8;\ s^2 = 4.$

O intervalo de confiança para a média é:

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$6,8 - (1,771)\frac{2}{\sqrt{14}} \le \mu \le 6,8 + (1,771)\frac{2}{\sqrt{14}}$$

$$\boxed{5,85336 \le \mu \le 7,74664}$$

φ \ α	0,40	0,25	0,10	0,05	0,025	0,010	0,005	0,0025	0,0010	0,0005
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657	127,321	318,309	636,619
2	0,289	0,816	1,886	2,920	4,303	6,975	9,925	14,089	22,327	31,599
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,215	12,924
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,327	3,787	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,659
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646

Intervalo de Confiança para a Variância

Para determinar uma distribuição que explique o comportamento da variância σ^2, considera-se o seguinte:

$$\frac{(X_i - \mu)}{\sigma} \sim N(0,1)$$

O quadrado dessa variável, tem distribuição qui-quadrado com 1 grau de liberdade.

$$\frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_1^2 \quad \Longrightarrow \quad \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_n^2$$

Já o somatório indicado acima, tem distribuição qui-quadrado com n graus de liberdade. Como μ é desconhecido, ele aproximado pela media amostral, sendo que a distribuição perde 1 grau de liberdade:

$$\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$$

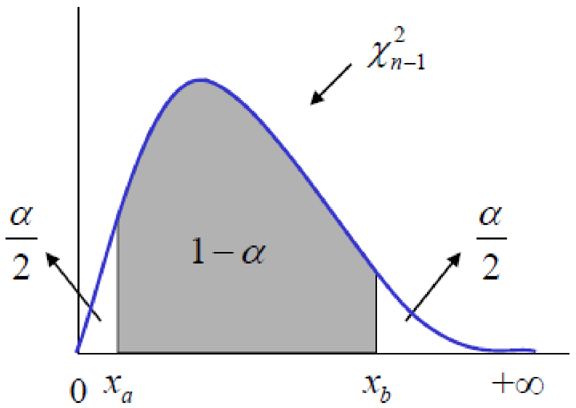
Pela variancia amostral, temos: $$\sum_{i=1}^{n} (X_i - \bar{X})^2 = s^2 (n-1)$$

Substituindo em termos da variancia amostral: $$\boxed{\frac{s^2 (n-1)}{\sigma^2} \sim \chi_{n-1}^2}$$

Intervalo de Confiança para a Variância

Como a variável χ^2_{n-1} tem distribuição qui-quadrado com (n-1) graus de liberdade:

$$\boxed{\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}}$$



Definem-se os limites x_a e x_b para um nível de confiança $(1-\alpha)$:

$$x_a \leq \frac{(n-1)s^2}{\sigma^2} \leq x_b$$

Intervalo de Confiança para a Variância

A variável χ^2_{n-1} tem distribuição qui-quadrado com (n-1) graus de liberdade, definem-se os limites x_a e x_b para um nível de confiança $(1-\alpha)$.

$$x_a \le \frac{(n-1)s^2}{\sigma^2} \le x_b$$

Escrevendo as desigualdades em função de σ^2, se tem: $\frac{(n-1)s^2}{x_b} \le \sigma^2 \quad \sigma^2 \le \frac{(n-1)s^2}{x_a}$

Assim, o intervalo de confiança para σ^2 é: $\boxed{\frac{(n-1)s^2}{x_b} \le \sigma^2 \le \frac{(n-1)s^2}{x_a}}$

Sendo que o nível de confiança $(1-\alpha)$, corresponde a probabilidade de que a variância σ^2 se encontre no intervalo calculado.

$$\boxed{P\left(\frac{(n-1)s^2}{x_b} \le \sigma^2 \le \frac{(n-1)s^2}{x_a}\right) = 1-\alpha}$$

Intervalo de Confiança para a Variância

Exemplo:
Uma variável aleatória contínua X tem distribuição desconhecida com média μ e variância σ^2. Retira-se uma amostra de 21 valores com variância $s^2 = 2{,}34$. Construir um intervalo de confiança de 95% para a variância populacional.

Resposta:
O intervalo de confiança para a variância é dado por:

$$\frac{(n-1)s^2}{x_b} \leq \sigma^2 \leq \frac{(n-1)s^2}{x_a}$$

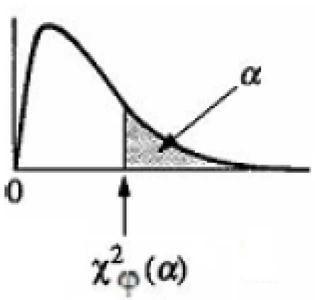
Sabe-e que n = 21 e $S^2 = 2{,}34$. Calcula-se os valores críticos x_a e x_b para a distribuição qui-quadrado com $(1-\alpha)=0{,}95$ e $\alpha=0{,}05$, sendo que:

$$x_a = \chi^2_{n-1}(1-\tfrac{\alpha}{2}) = \chi^2_{20}(1-0{,}025) = \chi^2_{20}(0{,}975)$$
$$x_b = \chi^2_{n-1}(\tfrac{\alpha}{2}) = \chi^2_{20}(0{,}025)$$

Os valores críticos são obtidos da tabela qui-quadrado.

Intervalo de Confiança para a Variância

Resposta: Os valores críticos são obtidos da tabela qui-quadrado:



$$x_a = \chi^2_{20}(0{,}975) = 9{,}59$$

$$x_b = \chi^2_{20}(0{,}025) = 34{,}17$$

O intervalo de confiança para a variância é:

$$\frac{(n-1)s^2}{x_b} \le \sigma^2 \le \frac{(n-1)s^2}{x_a}$$

$$\frac{(20)(2{,}34)}{34{,}17} \le \sigma^2 \le \frac{(20)(2{,}34)}{9{,}59}$$

$$\boxed{1{,}3696 \le \sigma^2 \le 4{,}8801}$$

φ	α 0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	0.000	0.000	0.001	0.004	0.015	0.102	0.455	1.323	2.71	3.84	5.02	6.63	7.88
2	0.010	0.020	0.050	0.103	0.211	0.575	1.386	2.77	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	1.213	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	1.923	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	2,67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00

Assim, existe 95% de chances da σ^2 se encontrar no intervalo.

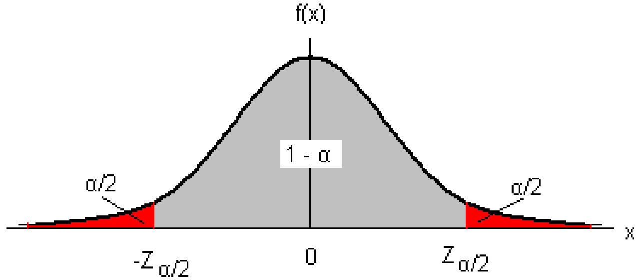
Intervalo de Confiança
para uma Proporção

Sabe-se que uma proporção tem distribuição normal:

$$\hat{p} \sim N\left(p, \frac{pq}{n}\right)$$

Padronizando a variável:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$$



Como a variável z tem distribuição normal, define-se os valores críticos $-Z_{\alpha/2}$ e $Z_{\alpha/2}$ para um nível de confiança (1-α):

$$-z_{\alpha/2} \leq \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \leq z_{\alpha/2}$$

Escrevendo as desigualdades em função de p:

$$\hat{p} + z_{\alpha/2}\sqrt{\frac{pq}{n}} \geq p \qquad p \geq \hat{p} - z_{\alpha/2}\sqrt{\frac{pq}{n}}$$

Se tem o intervalo para p:

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{pq}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{pq}{n}}$$

Como p é estimado por $\hat{p}$, então: $\sqrt{\frac{pq}{n}}$ é estimado por: $\sqrt{\frac{\hat{p}\hat{q}}{n}}$

Intervalo de Confiança
para uma Proporção

Assim, o intervalo de confiança para p é:

$$\boxed{\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

Sendo que o nível de confiança (1-α), corresponde a probabilidade de que a proporção p se encontre no intervalo calculado.

$$\boxed{P\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = 1 - \alpha}$$

Como o desvio padrão $\sqrt{pq}$

Intervalo de Confiança
para uma Proporção

Exemplo:
De 1220 peças fabricadas,1037 são perfeitas. Supondo o número de peças fabricadas está normalmente distribuído, construir um intervalo de confiança ao nível de1% de significância para a proporção de peças fabricadas que são defeituosas.

Resposta:
O intervalo de confiança para a proporção é dado por:

$$\boxed{\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

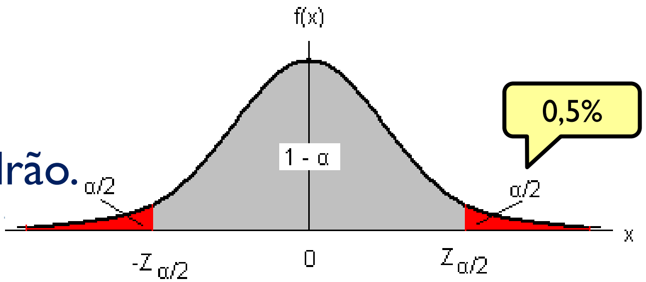
A proporção de peças defeituosas é dada por: $\frac{1220-1037}{1220} = \frac{183}{1220}$

Assim: $n = 1220;\ \hat{p} = \frac{183}{1220} = 0,15;\ \hat{q} = \frac{1037}{1220} = 0,85$

Calcula-se os valores críticos $-z_{\alpha/2}$ e $z_{\alpha/2}$ para a distribuição normal padrão com $(1-\alpha)=0,99$ e $\alpha=0,01$. Dada a simetria da normal somente é preciso calcular um valor crítico, sendo que:

$$z_{\alpha/2} = z_{0,01/2} = z_{0,005}$$

Os valores críticos são obtidos da tabela normal padrão.

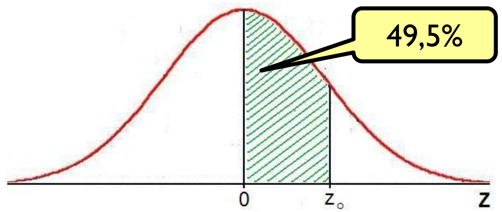


Intervalo de Confiança da Media

com Variância Conhecida

Intervalos de Confiança

Resposta: Da tabela normal.



$$P(0 < z < z_0) = \int_0^{z_0} \frac{1}{\sigma.\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$Z_{\alpha/2} = \quad Z_{0,495} = 2,57$$

Considerando: $n = 1220;\ \hat{\mathrm{p}} = 0,15;\ \hat{\mathrm{q}} = 0,85$

O intervalo de confiança para a proporção é:

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0,15 - (2,57)\sqrt{\frac{(0,15)(0,85)}{1220}} \le p \le 0,15 + (2,57)\sqrt{\frac{(0,15)(0,85)}{1220}}$$

$$\boxed{0,12372 \le p \le 0,17627}$$

z_0	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,49865	0,49869	0,49874	0,49878	0,49882	0,49886	0,49889	0,49893	0,49897	0,49900
3,1	0,49903	0,49906	0,49910	0,49913	0,49916	0,49918	0,49921	0,49924	0,49926	0,49929
3,2	0,49931	0,49934	0,49936	0,49938	0,49940	0,49942	0,49944	0,49946	0,49948	0,49950
3,3	0,49952	0,49953	0,49955	0,49957	0,49958	0,49960	0,49961	0,49962	0,49964	0,49965
3,4	0,49966	0,49968	0,49969	0,49970	0,49971	0,49972	0,49973	0,49974	0,49975	0,49976
3,5	0,49977	0,49978	0,49978	0,49979	0,49980	0,49981	0,49981	0,49982	0,49983	0,49983
3,6	0,49984	0,49985	0,49985	0,49986	0,49986	0,49987	0,49987	0,49988	0,49988	0,49989
3,7	0,49989	0,49990	0,49990	0,49990	0,49991	0,49991	0,49992	0,49992	0,49992	0,49992
3,8	0,49993	0,49993	0,49993	0,49994	0,49994	0,49994	0,49994	0,49995	0,49995	0,49995
3,9	0,49995	0,49995	0,49996	0,49996	0,49996	0,49996	0,49996	0,49996	0,49997	0,49997

Intervalo de Confiança
Diferença de duas Médias

Sabe-se que a diferença de duas médias tem distribuição normal:

$$(\bar{X}_1 - \bar{X}_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Padronizando a variável: $$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Como a variável z tem distribuição normal, define-se os valores críticos $-Z_{\alpha/2}$ e $Z_{\alpha/2}$ para um nível de confiança $(1-\alpha)$: $$-z_{\alpha/2} \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq z_{\alpha/2}$$

Escrevendo as desigualdades em função de $(\mu_1-\mu_2)$ se tem o intervalo de confiança:

$$\boxed{(\bar{X}_1 - \bar{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Quando as variâncias populacionais σ_1^2 e σ_2^2 são desconhecidos, utiliza-se as variancias amostrais s_1^2 e s_2^2 como estimadores.

Intervalo de Confiança
Diferença de duas Médias

Exemplo:

Foi escolhida uma turma de uma escola A e uma turma de uma escola B e aplicada uma prova de Probabilidade e Estatística. A turma da escola A com 50 alunos obteve média 76 com desvio padrão de 6 e a turma da escola B com 75 alunos obteve média 82 com desvio padrão de 8. Encontre um intervalo de confiança, com 5% de significância, para a diferença μ_1-μ_2, onde μ_1 representa a média de todos os alunos da escola A e μ2 a média de todos os alunos da escola B que poderiam fazer esta prova.

Intervalo de Confiança
Diferença de duas Proporções

Sabe-se que a diferença de duas proporções tem distribuição normal:

$$(\hat{p}_1 - \hat{p}_2) \sim N\left(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)$$

Padronizando a variável: $$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

Como a variável z tem distribuição normal, define-se os valores críticos $-Z_{\alpha/2}$ e $Z_{\alpha/2}$ para um nível de confiança $(1-\alpha)$:

$$-z_{\alpha/2} \leq \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \leq z_{\alpha/2}$$

Escrevendo as desigualdades em função de $(p_1 - p_2)$ se tem o intervalo:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Estimando-se p_1, p_2, q_1, q_2, se tem o intervalo de confiança:

$$\boxed{(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$$

Intervalo de Confiança
Diferença de duas Proporções

Exemplo:

Seja p_1 e p_2 as proporções de defeitos de um processo já existente e de um novo processo, respectivamente. Uma amostra de 1500 itens do processo já existente apresentou 75 itens com defeitos e uma amostra de 2000 itens do novo processo apresentou 80 itens com defeitos. Construir um intervalo de confiança de 90% para $p_1 - p_2$.