Seminário 3: Uso de estatística Z em uma amostra

Ex, Verificar aleatoriedade de amostra em relação a sobrevida de pacientes com câncer submetido a determinado tratamento versus a sobrevida normal dos pacientes

População: μ=38,3 meses σ=43,3 meses

Amostra: n=100, X=46,9

Estatística Z= $(\mu - x)/(\sigma/\sqrt{n})$

Z= $(46{,}9-38{,}3)/(43{,}3/\sqrt{100})$

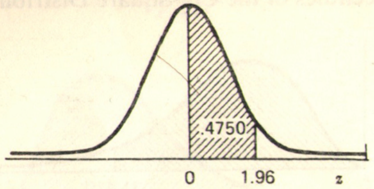
Z= 1,986

IC (5%): média amostral +/- (1,96 x EPM)

IC (5%): 46,9 (38,43-55,4)

Normal Curve Areas

Entries in the Body of the Table Give the Area Under the Standard Normal Curve from 0 to *z*.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

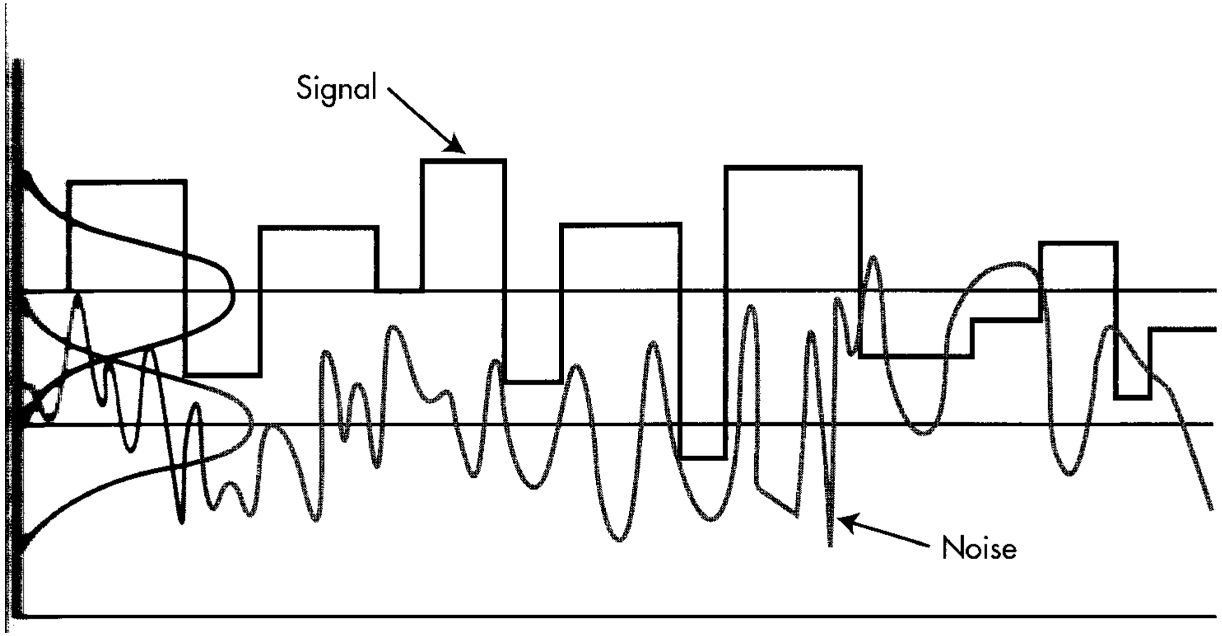
Área sob a curva normal com z=1,98: 0,4761

Ou seja, uma sobrevida de + 46,6 meses é alcançada por 0,4761 + 0,5 = 0,9761 (ou 97,61% da população)

Em outra palavras, a chance da média de sobrevida de uma amostra de 100 sujeitos 46,4 meses é de (1-0,9761=0,0239, ou 2,39%)

Teste t de Student: para amostras independentes

(não conhecemos nem a média nem o desvio padrão da população)



Suposições básicas de teste-t e ANOVA:

1. Variável contínua
2. Normalidade da população original
3. Homoscedasticidade (as variâncias entre as amostras <u>não são</u> diferentes) (mas se diferença entre o n da maior/menor amostra for menos que 1,5, usualmente não é importante)
4. Independência das amostras e das observações

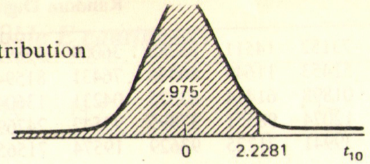
Teste t de Student: para amostras independentes

$$t = (x_1 - x_2)/\sqrt{}(S_1^2/n_1) + (S_2^2/n_2)$$

Graus de liberdade: número de pedaços únicos de informação em um conjunto de dados do qual conhecemos a soma. Em outras palavras, é o número de valores de um cálculo final de uma estatística que podem variar “livremente”.
No teste t= (n1 -1) + (n2 – 1)

Table H

Percentiles of the *t* Distribution



d. f.	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.9200	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409
4	1.533	2.1318	2.7764	3.747	4.6041
5	1.476	2.0150	2.5706	3.365	4.0321
6	1.440	1.9432	2.4469	3.143	3.7074
7	1.415	1.8946	2.3646	2.998	3.4995
8	1.397	1.8595	2.3060	2.896	3.3554
9	1.383	1.8331	2.2622	2.821	3.2498
10	1.372	1.8125	2.2281	2.764	3.1693
11	1.363	1.7959	2.2010	2.718	3.1058
12	1.356	1.7823	2.1788	2.681	3.0545
13	1.350	1.7709	2.1604	2.650	3.0123
14	1.345	1.7613	2.1448	2.624	2.9768
15	1.341	1.7530	2.1315	2.602	2.9467
16	1.337	1.7459	2.1199	2.583	2.9208
17	1.333	1.7396	2.1098	2.567	2.8982
18	1.330	1.7341	2.1009	2.552	2.8784
19	1.328	1.7291	2.0930	2.539	2.8609
20	1.325	1.7247	2.0860	2.528	2.8453
21	1.323	1.7207	2.0796	2.518	2.8314
22	1.321	1.7171	2.0739	2.508	2.8188
23	1.319	1.7139	2.0687	2.500	2.8073
24	1.318	1.7109	2.0639	2.492	2.7969
25	1.316	1.7081	2.0595	2.485	2.7874
26	1.315	1.7056	2.0555	2.479	2.7787
27	1.314	1.7033	2.0518	2.473	2.7707
28	1.313	1.7011	2.0484	2.467	2.7633
29	1.311	1.6991	2.0452	2.462	2.7564
30	1.310	1.6973	2.0423	2.457	2.7500
35	1.3062	1.6896	2.0301	2.438	2.7239
40	1.3031	1.6839	2.0211	2.423	2.7045
45	1.3007	1.6794	2.0141	2.412	2.6896
50	1.2987	1.6759	2.0086	2.403	2.6778
60	1.2959	1.6707	2.0003	2.390	2.6603
70	1.2938	1.6669	1.9945	2.381	2.6480
80	1.2922	1.6641	1.9901	2.374	2.6388
90	1.2910	1.6620	1.9867	2.368	2.6316
100	1.2901	1.6602	1.9840	2.364	2.6260
120	1.2887	1.6577	1.9799	2.358	2.6175
140	1.2876	1.6558	1.9771	2.353	2.6114
160	1.2869	1.6545	1.9749	2.350	2.6070
180	1.2863	1.6534	1.9733	2.347	2.6035
200	1.2858	1.6525	1.9719	2.345	2.6006
∞	1.282	1.645	1.96	2.326	2.576

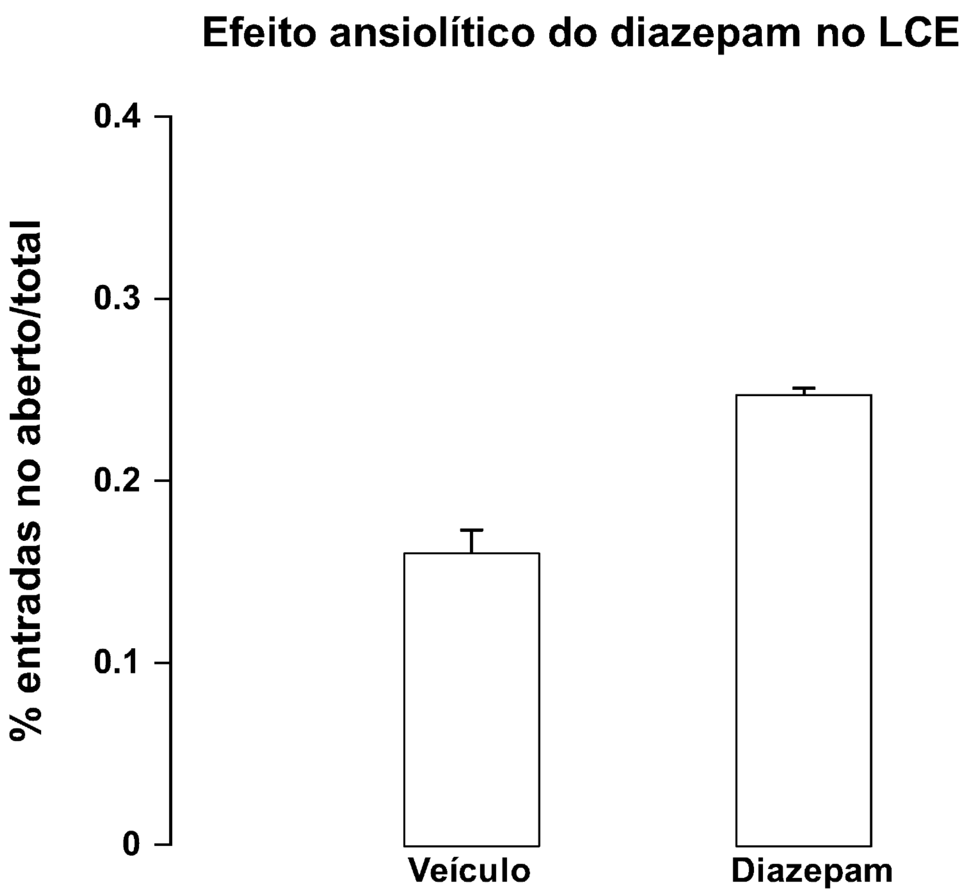
Distribuição t: média=0, simétrica, é uma família de distribuições (diferente para cada g.l., mas se aproxima da normal quando n-1 se aproxima do infinito)

Veículo	Diazepam (2mg/kg)
,15	0,20
,17	0,24
,18	0,15
,10	0,28
,09	0,32
,18	0,30
,19	0,38
,12	0,4
	0,2
	0,15
	0,10

Group Statistics

	droga	N	Mean	Std. Deviation	Std. Error Mean
Percab	sal	8	,1475	,03919	,01386
	diazepam 2 mg;kg	11	,2491	,12950	,03904

Serão diferente?


Efeito ansiolítico do diazepam no LCE
% entradas no aberto/total
0.4
0.3
0.2
0.1
0
Veículo
Diazepam

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Percab	Equal variances assumed	12,630	,002	-2,134	17	,048	-,10159	,04761	-,20203	-,00115
	Equal variances not assumed			-2,452	12,396	,030	-,10159	,04143	-,19154	-,01164

"Effect size"

ETA-squared (representado por η^2): expressa a % da variância da variável explicada pelo fator. Seu cálculo é fácil:

Eta-squared= $t^2/[t^2 + (N1 + N2 - 2)]$

Um η^2 de 0.01 seria considerado pequeno, um de 0.06 moderado e > de 0.14 grande (Stevens J. Applied multivariade statistics for the social sciences. 4th Ed., Lawrance Erlbaum Associates, Mahwah, NJ, 2001).

η^2=4,41/(64,41+ 17)= 0,20

Table H

Percentiles of the *t* Distribution

d. f.	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.9200	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409
4	1.533	2.1318	2.7764	3.747	4.6041
5	1.476	2.0150	2.5706	3.365	4.0321
6	1.440	1.9432	2.4469	3.143	3.7074
7	1.415	1.8946	2.3646	2.998	3.4995
8	1.397	1.8595	2.3060	2.896	3.3554
9	1.383	1.8331	2.2622	2.821	3.2498
10	1.372	1.8125	2.2281	2.764	3.1693
11	1.363	1.7959	2.2010	2.718	3.1058
12	1.356	1.7823	2.1788	2.681	3.0545
13	1.350	1.7709	2.1604	2.650	3.0123
14	1.345	1.7613	2.1448	2.624	2.9768
15	1.341	1.7530	2.1315	2.602	2.9467
16	1.337	1.7459	2.1199	2.583	2.9208
17	1.333	1.7396	2.1098	2.567	2.8982
18	1.330	1.7341	2.1009	2.552	2.8784
19	1.328	1.7291	2.0930	2.539	2.8609
20	1.325	1.7247	2.0860	2.528	2.8453
21	1.323	1.7207	2.0796	2.518	2.8314
22	1.321	1.7171	2.0739	2.508	2.8188
23	1.319	1.7139	2.0687	2.500	2.8073
24	1.318	1.7109	2.0639	2.492	2.7969
25	1.316	1.7081	2.0595	2.485	2.7874
26	1.315	1.7056	2.0555	2.479	2.7787
27	1.314	1.7033	2.0518	2.473	2.7707
28	1.313	1.7011	2.0484	2.467	2.7633
29	1.311	1.6991	2.0452	2.462	2.7564
30	1.310	1.6973	2.0423	2.457	2.7500

Effect size (Cohen's d)

d= (média1 – média2)/DP (pooled)

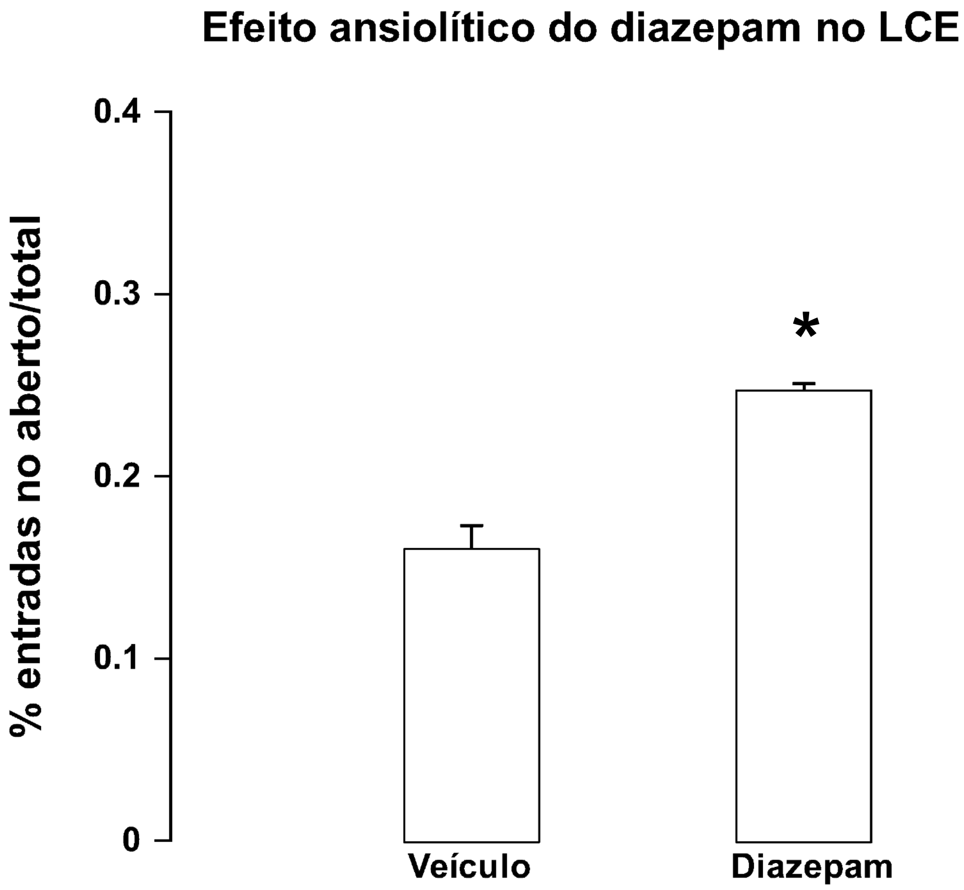
Segundo Cohen: 0,2= pequeno
0,5= moderado
>0,8=grande

Pode calcuar só com o valor dos n e do t

$d = t \times \sqrt{(1/n_1) + (1/n_2)}$

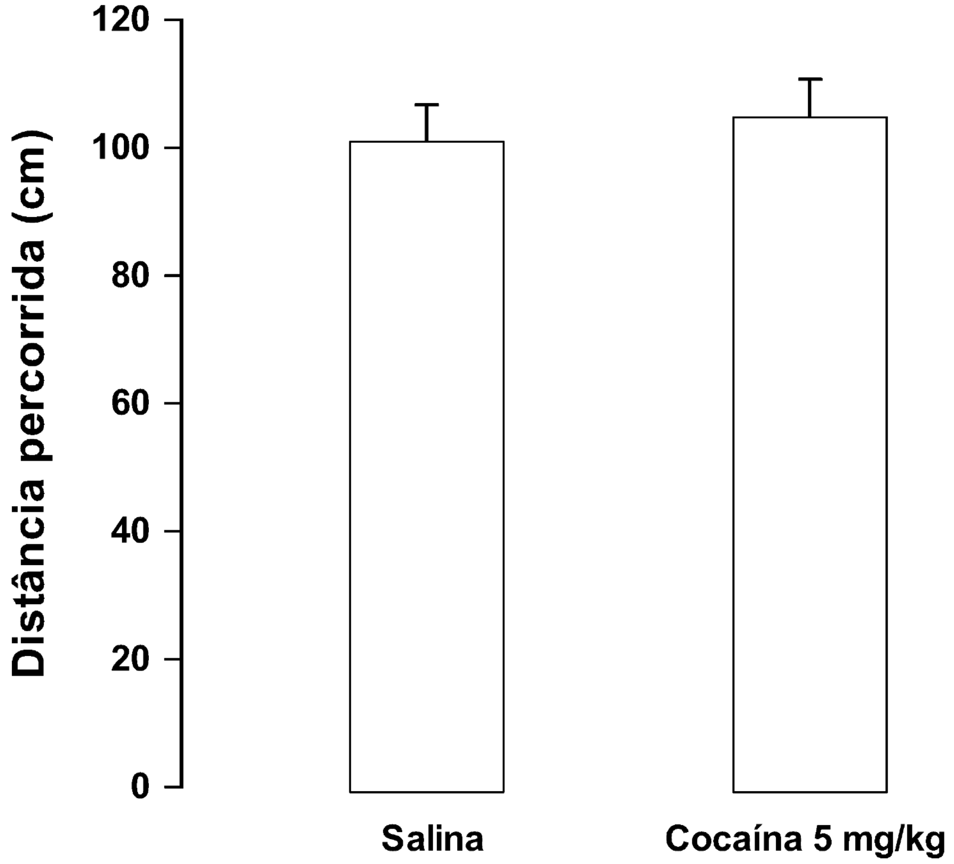
No caso $d = 2{,}13 \times \sqrt{(1/7) + (1/11)}$

d = 1,029


Efeito ansiolítico do diazepam no LCE
% entradas no aberto/total
0.4
0.3
0.2
0.1
0
*
Veículo
Diazepam

Teste t-pareado

Salina	Cocaína
120,00	125,00
130,00	131,00
110,00	120,00
80,00	85,00
90,00	95,00
71,00	74,00
114,00	121,00
100,00	101,00
98,00	99,00
97,00	97,00



Distância percorrida (cm)
120
100
80
60
40
20
0
Salina
Cocaína 5 mg/kg

t-tests for independent samples of TRAT

Variable	Number of Cases	Mean	SD	SE of Mean
DISTPERC				
TRAT 1,	10	101,0000	18,074	5,715
TRAT 2,	10	104,8000	18,660	5,901

Mean Difference = -3,8000

Levene's Test for Equality of Variances: F= ,129 P= ,723

t-test for Equality of Means

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-,46	18	,649	8,215	(-21,063; 13,463)
Unequal	-,46	17,98	,649	8,215	(-21,063; 13,463)

Teste t de Student: para amostras repetidas

t= diferença entre médias/EP das diferenças

Teste t-pareado

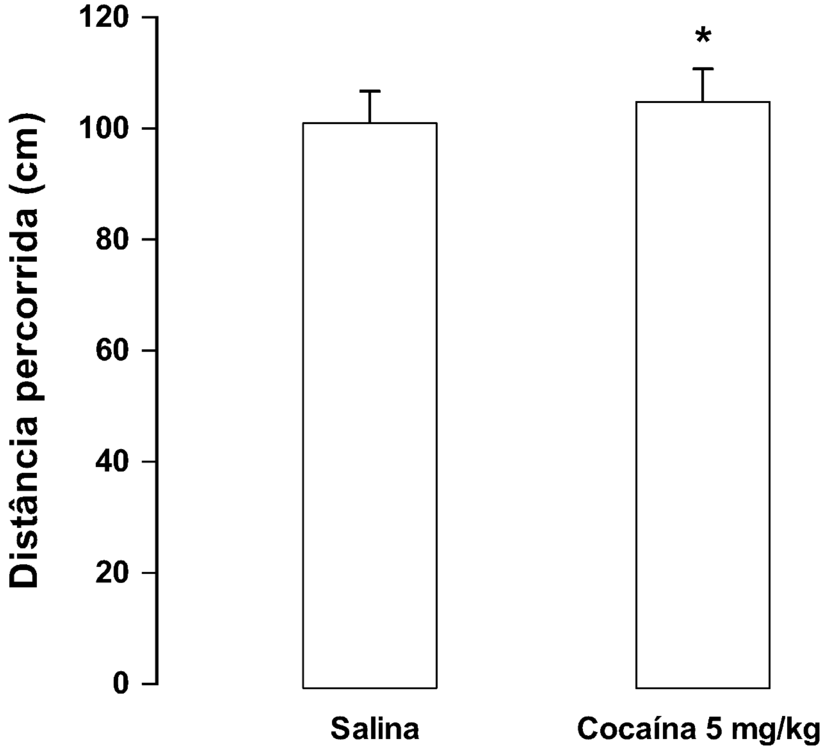
Salina	Cocaína	Diferença	$(n-X)^2$
120,00	125,00	5	1,44
130,00	131,00	1	7,84
110,00	120,00	10	38,44
80,00	85,00	5	1,44
90,00	95,00	5	1,44
71,00	74,00	3	0,64
114,00	121,00	7	10,24
100,00	101,00	1	7,84
98,00	99,00	1	7,84
97,00	97,00	0	14,44
Média (SEM)		3,8 (1,0)	

S^2=10,17
S=3,19
EP=1,009

- - - t-tests for paired samples - - -

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
DISTDRUG	10	,985	,000	104,8000	18,660	5,901
DISTSAL				101,0000	18,074	5,715

Paired Differences					
Mean	SD	SE of Mean	t-value	df	2-tail Sig
3,8000	3,190	1,009	3,77	9	,004
95% CI (1,517; 6,083)					


Distância percorrida (cm)
120
100
80
60
40
20
0
*
Salina
Cocaína 5 mg/kg