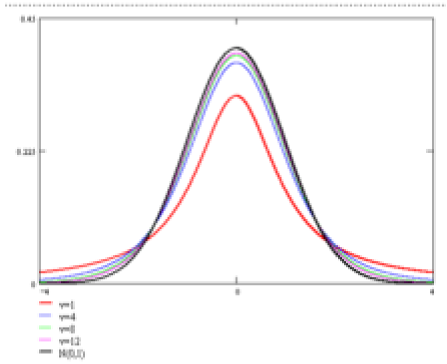
WIKIPÉDIA

Distribuição t de Student

Origem: Wikipédia, a enciclopédia livre.

A **distribuição *t* de Student** é uma distribuição de probabilidade estatística, publicada por um autor que se chamou de *Student*, pseudônimo de William Sealy Gosset, que não podia usar seu nome verdadeiro para publicar trabalhos enquanto trabalhasse para a cervejaria Guinness.[1][2]



A função densidade da distribuição de Student para alguns valores de *v* e da distribuição normal (a preto).

A distribuição t é uma distribuição de probabilidade teórica. É simétrica, campaniforme, e semelhante à curva normal padrão, porém com caudas mais largas, ou seja, uma simulação da t de Student pode gerar valores mais extremos que uma simulação da normal. O único parâmetro *v* que a define e caracteriza a sua forma é o número de *graus de liberdade*. Quanto maior for esse parâmetro, mais próxima da normal ela será.

Índice

Definição

Suponha *Z*, uma variável aleatória de distribuição normal padrão com média 0 e variância 1, e *V*, uma variável aleatória com distribuição Chi-quadrado com *v* graus de liberdade. Se *Z* e *V* são independentes, então a distribuição da variável aleatória t será[3]:

$$t = \frac{Z}{\sqrt{V/\nu}}$$

Essa é a **distribuição *t* de Student** com *v* graus de liberdade.

Função densidade de probabilidade

A função densidade de probabilidade é:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-(\frac{\nu+1}{2})},$$

em que Γ é a função gama. Usando-se a função beta *B*, a função densidade de probabilidade pode ser escrita como:

$$f(t) = \frac{1}{\sqrt{\nu}\,B\left(\frac{1}{2},\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-(\frac{\nu+1}{2})},$$

Aplicações

A distribuição t de Student aparece naturalmente no problema de se determinar a média de uma população (que segue a distribuição normal) a partir de uma amostra. Neste problema, não se sabe qual é a média ou o desvio padrão da população, mas ela deve ser normal.

Supondo que o tamanho da amostra *n* seja muito menor que o tamanho da população, temos que a amostra é dada por *n* variáveis aleatórias normais independentes X_1, ..., X_n, cuja média $\overline{X}_n = (X_1 + \cdots + X_n)/n$ é o melhor estimador para a média da população.

Considerando ${S_n}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$ como a variância amostral, temos o seguinte resultado:

A variável aleatória *t* dada por:

$$t = \frac{\overline{X}_n - \mu}{S_n/\sqrt{\nu}},$$

ou :$t = \sqrt{\nu}\frac{\overline{X}_n - \mu}{S_n}$, segue uma distribuição t de Student com $\nu = n - 1$ graus de liberdade.

Tabela com alguns valores selecionados

Grande parte dos livros estatísticos trazem uma tabela com valores para a distribuição t de Student. Essas tabelas apresentam valores arredondados e esses arredondamentos podem ser grosseiros demais, dependendo do tipo de análise que está sendo feita. Softwares estatísticos e planilhas como Microsoft Excel e OpenOffice Calc possuem técnicas mais precisas para a estimação desses valores.

A tabela abaixo lista alguns valores selecionados para a distribuição t de Student com ν graus de liberdade (números no início de cada linha) para as regiões críticas com uma ou duas caudas (unicaudal ou bicaudal). Por exemplo, se estamos fazendo uma análise em que a distribuição t de Student apresenta 4 graus de liberdade e queremos usar um nível de confiança de 95% unicaudal, consultamos a tabela e percebemos que t deve ser de 2,132. Isso quer dizer que a probabilidade de $-\infty < t < 2,132$ é de 95%.

Unicaudal	**75%**	**80%**	**85%**	**90%**	**95%**	**97,5%**	**99%**	**99,5%**	**99,75%**	**99,9%**	**99,95%**
Bicaudal	**50%**	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99,5%**	**99,8%**	**99,9%**
1	1,000	1,376	1,963	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,767
24	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,659
30	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
50	0,679	0,849	1,047	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
60	0,679	0,848	1,045	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460
80	0,678	0,846	1,043	1,292	1,664	1,990	2,374	2,639	2,887	3,195	3,416
100	0,677	0,845	1,042	1,290	1,660	1,984	2,364	2,626	2,871	3,174	3,390
120	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	2,860	3,160	3,373

∞	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Exemplo

Um fabricante de aparelhos celulares afirma que a duração média de sua bateria nos primeiros 6 meses de uso é de 120 horas,ou seja, 5 dias. Analisando uma amostra de 25 aparelhos, obteve-se uma média de duração de 116 horas, com desvio padrão de 12 horas. Verifique se a afirmação é verdadeira, utilizando um nível de confiança de 95% bicaudal.

Resolução:

1º Utilizando a tabela de distribuição t student, definem-se os pontos críticos através do grau de liberdade (24) e o nível de confiança (95%).

Nesse caso, os pontos críticos são **± 2,064**, ou seja, **P(-2,064 < t < 2,064)**. Se o valor de t estiver dentro desses limites a afirmação é verdadeira.

2º Na sequência calcula-se o valor de t para a amostra:

Dados:

$\overline{X} = 116;$

$\mu = 120;$

$S = 12;$

$n = 25;$

Fórmula: $t = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{116 - 120}{12/\sqrt{25}} = -1,667$

3º Conclusão: Como **t = - 1,667**, encontra-se dentro dos limites críticos, **P(-2,064 < t < 2,064)**, a afirmação do fabricante de celular que a duração média da sua bateria é de 120 horas, a um nível de confiança de 95%, é verdadeira.

Ver também

- Distribuição normal

Referências

1. William Gosset (http://alea-estp.ine.pt/html/nomesedatas/swf/biografias.asp?art=9), *site* da Acção Local Estatística Aplicada (http://alea.ine.pt)
2. História da Estatística no mundo (http://www.estatistica.ccet.ufrn.br/historia.php), *site* da UFRN
3. Myers, Raymond H. (2009). *Probabilidade e estatística para engenharia e ciências (8a Edição)*. São Paulo: Pearson Education do Brasil. pp. 162–163. ISBN 978-85-430-1440-1