11 Confidence Interval for a Correlation Coefficient

A research team conducted a longitudinal study of participants between 25 and 30 years of age. They measured each participant's level of education at age 25. They also measured each participant's earnings at age 30. The team collected data on $n = 314$ participants. The average level of education at age 25 was 15.3, with an observed standard deviation of 3.1 (the divisor in the underlying variance calculation was $n - 1$). The average earnings (in thousands of dollars) was 54.9, with an observed standard deviation of 14.9 (the divisor in the underlying variance calculation was $n-1$). The Pearson product moment correlation coefficient between the two variables was 0.76. The research team seeks to estimate the regression of participant earnings at age 30 on participant education at age 25.

e. What is the 95% confidence interval for the population correlation coefficient of level of education at age 25 and earnings at age 30.

USE $F(\rho) = \frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right)$

$F(r) = F(0.76) = \frac{1}{2} \ln\left(\frac{1.76}{0.24}\right) = 0.9962$

WHEN R IS PEARSON PRODUCT MOMENT CORRELATION,

$F(R) \sim N\left(\frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right), \frac{1}{n-3}\right).$

HENCE 95% SAMPLING MARGIN OF ERROR IS

$1.960\sqrt{\frac{1}{n-3}} = 1.960\sqrt{\frac{1}{311}} = 0.1111$

95% CI FOR $F(\rho)$ IS

0.9962 ± 0.1111

$= .8851$ TO 1.107

BUT NEED CI FOR ρ.

USE INVERSION FORMULA.

FOR $F(\rho) = \frac{1}{2} \ln\left(\frac{1+\rho}{1-\rho}\right)$,

$$\rho = \frac{\left[e^{2F(\rho)}\right] - 1}{\left[e^{2F(\rho)}\right] + 1}.$$

FOR LEFT ENDPOINT 0.8851,

$$e^{2(0.8851)} = 5.872$$

$$\rho_{LEFT} = \frac{5.872 - 1}{5.872 + 1} = 0.709$$

FOR RIGHT ENDPOINT 1.107,

$$e^{2(1.107)} = 9.152$$

$$\rho_{RIGHT} = \frac{9.152 - 1}{9.152 + 1} = 0.803$$

WHEN $r = 0.76$, $n = 314$, THE

95% CONFIDENCE INTERVAL

FOR ρ IS

$$\boxed{0.709 \text{ TO } 0.803}$$

3

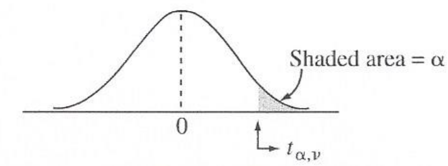

TABLE 2

Percentage points of Student's t distribution

	Right-Tail Probability (α)								
df	**.40**	**.25**	**.10**	**.05**	**.025**	**.01**	**.005**	**.001**	**.0005**
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	.289	.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	.277	.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
35	.255	.682	1.306	1.690	2.030	2.438	2.724	3.340	3.591
40	.255	.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	.255	.679	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	.254	.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
inf.	.253	.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: Computed by M. Longnecker using the R function $qt(1 - \alpha, \text{df})$.

For level α two-tailed tests and $100(1 - \alpha)\%$ C.I.s use value in column headed by the number obtained by computing $\alpha/2$.