

ECON 6511: Econometrics
Practice Final Exam
Points: 70, Time: 210 minutes

Instructions

- Use 3 or more decimal places unless otherwise stated
- No notes or cellphones

Section A: Multiple-choice (15 points)

1. If an explanatory variable is strictly exogenous it implies that:

- (a) changes in the lag of the variable do not affect future values of the dependent variable.
- (b) the variable is correlated with the error term in all future time periods.
- (c) the variable cannot react to what has happened to the dependent variable in the past.
- (d) the explanatory variable is not contemporaneously exogenous.

Answer: (c)

2. A covariance stationary time series is weakly dependent if:

- (a) the correlation between the independent variable at time t and the dependent variable at time $t + h$ goes to ∞ as $h \rightarrow 0$.
- (b) the correlation between the independent variable at time t and the dependent variable at time $t + h$ goes to 0 as $h \rightarrow \infty$.
- (c) the correlation between the independent variable at time t and the independent variable at time $t + h$ goes to ∞ as $h \rightarrow 0$.
- (d) the correlation between the independent variable at time t and the independent variable at time $t + h$ goes to 0 as $h \rightarrow \infty$.

Answer: (d)

3. A pooled OLS estimator that is based on the time-demeaned variables is called the ...

- (a) random effects estimator
- (b) fixed effects estimator
- (c) least absolute deviations estimator
- (d) instrumental variable estimator

Answer: (b)

4. What should be the degrees of freedom (df) for fixed effects estimation if the data set includes N cross sectional units over T time periods and the regression model has k independent variables?

- (a) $N - kT$
- (b) $NT - k$
- (c) $NT - N - k$
- (d) $N - T - k$

Answer: (c)

5. Which of the following is a property of dummy variable regression?

- (a) This method is best suited for panel data sets with many cross-sectional observations.
- (b) The R-squared obtained from this method is lower than that obtained from regression on time-demeaned data.
- (c) The degrees of freedom cannot be computed directly with this method.
- (d) The major statistics obtained from this method are identical to that obtained from regression on time-demeaned data.

Answer: (d)

6. The estimator obtained through regression on quasi-demeaned data is called the

- (a) random effects estimator
- (b) fixed effects estimator
- (c) hetroskedasticity-robust OLS estimator
- (d) instrumental variables estimator

Answer: (a)

7. An economist wants to study the effect of income on savings. He collected data on 120 identical twins. Which of the following methods of estimation is the most suitable method, if income is correlated with the unobserved family effect?

- (a) Random effects estimation
- (b) Fixed effects estimation
- (c) Ordinary least squares estimation
- (d) Weighted Least squares estimation

Answer: (b)

8. Consider the following simple regression model: $y = \beta_0 + \beta_1 x_1 + u$. In order to obtain consistent estimators of β_0 and β_1 , when x and u are correlated, a new variable z is introduced into the model which satisfies the following two conditions: $Cov(z, x) \neq 0$ and $Cov(z, u) = 0$. The variable z is called a(n) ... variable.

- (a) dummy
- (b) instrumental
- (c) lagged dependent variable
- (d) random

Answer: (b)

9. Consider the following simple regression model $y = \beta_0 + \beta_1 x_1 + u$. Suppose z is an instrument for x . if $Cov(z, u) = 0$ and $Cov(z, x) \neq 0$, the value of β_1 in terms of population covariances is

- (a) $(Cov(z, y))/(Cov(z, x))$
- (b) $(Cov(z, u))/(Cov(z, x))$
- (c) $Cov(z, u)$
- (d) $Cov(z, x)$

Answer: (a)

10. The sampling variance for the instrumental variables (IV) estimator is larger than the variance for the ordinary least square estimators (OLS) because

- (a) $R^2 > 1$
- (b) $R^2 < 0$
- (c) $R^2 = 1$
- (d) $R^2 \leq 1$

Answer: (d)

11. Consider the following simple regression model $y = \beta_0 + \beta_1 x_1 + u$. The variable z is a poor instrument for x if

- (a) there is a high correlation between z and x
- (b) there is a low correlation between z and x

- (c) there is a high correlation between z and u
- (d) there is a low correlation between z and u

Answer: (b)

12. Which of the following correctly identifies a characteristic of structural equations?
- (a) A structural equation should contain equal number of dependent and independent variables.
 - (b) A structural equation should contain equal number of endogenous and exogenous variables.
 - (c) A structural equation should have a behavioral, ceteris paribus interpretation on its own.
 - (d) A structural equation should not contain structural errors.

Answer: (c)

13. The model: $G(z) = [exp(z)]/[1+exp(z)]$, where G is between zero and one for all real numbers z , represents a:
- (a) logit model.
 - (b) probit model.
 - (c) Tobit model.
 - (d) linear probability model.

Answer: (a)

14. The model: $G(z) = \int_{-\infty}^z \phi(v)dv$ where $\phi(z)$ denotes the standard normal pdf represents a:
- (a) Tobit model.
 - (b) logit model.
 - (c) probit model.
 - (d) linear probability model.

Answer: (c)

15. A count variable refers to a dependent variable that can take on:
- (a) nonnegative integer values.
 - (b) nonnegative fractional values.
 - (c) negative fractional values.
 - (d) negative integer values.

Answer: (a)

Section B: Written Answer (55 points)

1. **(15 points)** Consider the following model of scores on a standardized math exam in 550 districts between 1993 and 1998:

$$\begin{aligned} \text{math4}_{it} = & \beta_0 + \beta_1 \log(\text{rexpp}_{it}) + \beta_2 \log(\text{rexpp}_{i,t-1}) + \beta_3 \log(\text{enrol}_{it}) + \beta_4 \text{lunch} \\ & + \beta_5 y94 + \dots + \beta_9 y98 + a_i + u_{it} \end{aligned}$$

where *math4* is the percentage of fourth graders in a district receiving a passing score on a standardized math test, *rexpp* is the value of real expenditures per student in the district, *enrol* is total district enrollment, *lunch* is the percentage of students eligible for the school lunch program (which is determined by poverty status), and *y94* is a dummy variable for 1994 (1993 is the excluded year). Results by pooled OLS, random effects, and fixed effects are given below.

- (a) Using pooled OLS, what is the interpretation of the 1995 dummy variable?

Answer: The percentage of satisfactory scores is 18.65 percent higher in 1995 compared to 1993.

- (b) Using pooled OLS, what is the interpretation of the coefficient on $\log(\text{rexpp}_{it})$?

Answer: Each 1% increase in expenditure this year increases scores by $\frac{0.534}{100} = 0.005$ percent.

- (c) Using pooled OLS, what is the interpretation of the coefficient on *lunch*?

Answer: A one percentage point increase in *lunch* reduces the pass rate by 0.41 percentage points.

- (d) Suppose the residuals from estimating the pooled OLS regression, \hat{v}_{it} were regressed on the lag of these residuals, $\hat{v}_{i,t-1}$. If the coefficient on $\hat{v}_{i,t-1}$ is $\hat{\rho} = 0.504$ and the standard error is $se = 0.017$, what does this imply?

Answer: This implies there is very strong evidence of (AR(1)) serial correlation, which is not surprising since there are time-constant unobserved variables in each district, a_i , meaning $Cov(a_i + u_{it}, a_i + u_{i,t+1}) \neq 0$.

- (e) Based on your answer to the above question, are the standard errors for the pooled OLS estimates likely to be valid? Explain.

Answer: Since there are likely to be unobserved factors contained in a_i , there will be serial correlation so the standard errors will not be valid.

- (f) List one factor that might be contained in a_i and explain how it might be correlated with *lunch*.

Answer: Average income is an obvious omitted variable. It is likely to be strongly

(negatively) correlated with *lunch* and positively correlated with *math4* so the coefficient on *lunch* will be biased downwards.

- (g) What happens to the coefficient on *lunch* when the model is estimated using fixed effects? Explain.

Answer: The coefficient rises from -0.407 to 0.062 . Assuming that *lunch* is slow to change over time, their effects are mostly explained by the unobserved effect, a_i . Estimation using fixed effects seems to have removed some of the downward bias suggested in the previous question.

- (h) Suppose you estimated the model using OLS but included dummy variables for each district as opposed to time demeaning or quasi demeaning the data. What would you expect to get for your estimate of β_4 ?

Answer: This will give identical coefficients and standard errors to the fixed effects regression so the estimated coefficient on *lunch* will be 0.062 .

- (i) What is your estimate of the long-run spending effect $\theta = \beta_1 + \beta_2$ using the fixed effects model? Explain how you would estimate its standard error?

Answer: The estimate is $\hat{\theta} = -0.411 + 7.00 = 6.59$. To obtain its standard error, we would estimate the following regression, obtained by substituting in $\theta - \beta_2$ for β_1 :

$$\begin{aligned} \text{math4}_{it} = & \beta_0 + \theta \log(\text{rexp}_{it}) + \beta_2 (\log(\text{rexp}_{i,t-1}) - \log(\text{rexp}_{it})) \\ & + \beta_3 \log(\text{enrol}_{it}) + \beta_4 \text{lunch} + \beta_5 y94 + \dots + \beta_9 y98 + a_i + u_{it} \end{aligned}$$

- (j) Suppose that you estimated the model using random effects and the quasi-demeaning parameter is 0.58 . What does this imply?

Answer: It means that the random effects and fixed effects estimates are not very similar. If the parameter was close to 1 then they would be. You can see that the estimates are quite different, e.g. the coefficient on *lunch* is -0.334 under random effects but 0.062 under fixed effects.

	Pooled OLS	Fixed Effects	Random Effects
log(rexpp)	0.534 (2.428)	-0.411 (2.458)	0.382 (2.060)
log(rexpp)_1	9.049** (2.305)	7.003** (2.369)	7.806** (1.925)
log(enrol)	0.593** (0.205)	0.245 (1.100)	0.787* (0.347)
lunch	-0.407** (0.014)	0.062 (0.051)	-0.334** (0.023)
y94	6.377** (0.736)	6.177** (0.560)	6.357** (0.560)
y95	18.650** (0.786)	18.093** (0.691)	18.642** (0.630)
y96	18.033** (0.767)	17.940** (0.757)	18.204** (0.651)
y97	15.340** (0.777)	15.192** (0.799)	15.518** (0.673)
y98	30.398** (0.783)	29.883** (0.837)	30.543** (0.691)
Constant	-31.662** (10.301)	-16.081 (23.807)	-23.224 (14.752)
Observations	3300	3300	3300
R^2	0.505	0.603	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

2. (15 points) Consider the following model of female labor force participation:

$$inlf_i = \beta_0 + \beta_1 educ_i + \beta_3 exper_i + \beta_4 exper_i^2 + \beta_5 age_i + u_i$$

where $inlf$ is an indicator denoting whether i is in the labor force. Linear probability model, probit, and logit results are given below.

- (a) Would it make sense to estimate this model using a poisson regression as well? Explain.
Answer: No, the left-hand side variable is binary. Poisson models are more appropriate for left-hand side variables that take on non-negative integer values.
- (b) Interpret the estimated coefficient $\hat{\beta}_1$ using the LPM results.
Answer: One extra year of schooling increases the probability of being in the labor force by 2.8 percent.
- (c) Using the LPM estimates, what is the predicted probability of a 35 year-old woman being in labor force for which $educ = 12$ and $exper = 10$?
Answer: The predicted probability is: $0.028 \cdot 12 + 0.045 \cdot 10 - 0.001 \cdot 100 - 0.011 \cdot 35 + 0.327 = 0.336 + 0.45 - 0.1 = 0.628$ or 62.8%.
- (d) Why are the logit estimates larger than the probit estimates?
Answer: Logit estimates are typically $\frac{0.4}{0.25} = 1.6$ times larger since $g(0) = 0.4$ for the probit model and $g(0) = 0.25$ for the logit model.
- (e) Explain how you would test whether the variables $exper$ and $exper^2$ are jointly significant in the logit model?
Answer: Conduct a likelihood ratio test. Run a logit regression that excludes $exper$ and $exper^2$. This is the restricted model. The unrestricted model is given in column 3. The test statistic is: $LR = 2(\log L_{ur} - \log L_r) \sim \chi_q^2$, where $q = 2$ is the degrees of freedom (number of restrictions) and L is the (log) likelihood for each model.
- (f) Explain how you would calculate the same predicted probability from part (c) using the logit estimates.
Answer: The probability would be equal to $\Lambda(0.144 \cdot 12 + 0.211 \cdot 10 - 0.003 \cdot 100 - 0.052 \cdot 35 - 0.882)$ where $\Lambda(z) = \frac{1}{1+e^{-z}}$ is equal to the logistic cumulative distribution function.
- (g) Explain how you would calculate the marginal effect at the average for $educ$ (not the average marginal effect) for either the logit or probit model.
Answer: This is equal to $g(0.144 \cdot \overline{educ} + 0.211 \cdot \overline{exper} - 0.003 \cdot \overline{exper^2} - 0.052 \cdot \overline{age} - 0.882) \times 0.144$, where the overlines denote the average values in the sample.

	LPM	Probit	Logit
<i>educ</i>	0.028** (0.007)	0.086** (0.022)	0.144** (0.038)
<i>exper</i>	0.045** (0.006)	0.129** (0.018)	0.211** (0.031)
<i>exper</i> ²	-0.001** (0.000)	-0.002** (0.001)	-0.003** (0.001)
<i>age</i>	-0.011** (0.002)	-0.032** (0.007)	-0.052** (0.011)
Constant	0.327* (0.137)	-0.521 (0.413)	-0.882 (0.686)
Observations	753	753	753
<i>R</i> ²	0.194		
log-likelihood		-401.30219	-401.76515

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

3. **(15 points)** A common method for estimating Engel curves is to model expenditure shares as a function of total expenditure, and possibly demographic variables. An Engel curve describes how household expenditure on a particular good or service varies with household income. A common specification has the form

$$sgood = \beta_0 + \beta_1 \log(totexpend) + demographics + u,$$

where *sgood* is the fraction of spending on a particular good out of total expenditure and *totexpend* is the log of total expenditure. The sign and magnitude of β_1 are of interest across various expenditure categories. To account for the potential endogeneity of *totexpend* – which can be viewed as an omitted variables or simultaneous equations problem, or both – the log of family income is often used as an instrumental variable. Let *lincome* denote the log of family income. Various estimates are provided below.

- (a) In the first model, *sfood*, the share of spending (0 to 1) on food, is used as the dependent variable. Would you expect there to be many zeros for *sfood*?

Answer: No, the necessity of food would make zeros unlikely (though there could be some families that are completely self-sufficient).

- (b) Interpret the estimated coefficient on $\log(totexpend)$ from the first model.

Answer: If $\log(\text{totexpend})$ increases by 0.1, which means a 10% increase in total expenditure, the share of food falls by 0.015 or 1.5 percentage points.

- (c) Why might $\log(\text{totexpend})$ be endogenous? Suggest one possibility.

Answer: u might contain a race variable such as *white* and this might be correlated with ltotexpend (perhaps white people spend more/less than black or hispanic people).

- (d) What evidence is there that $\log(\text{income})$ is a valid IV for $\log(\text{totexpend})$? Explain.

Answer: Column 2 indicates that $\log(\text{income})$ is strongly correlated with $\log(\text{totexpend})$.

- (e) What else do we require for $\log(\text{income})$ to be an instrument? Can this be tested also?

Answer: We require that $\log(\text{income})$ is not correlated with u . Unfortunately, this cannot be tested.

- (f) How does the IV estimate of β_1 compare to the OLS estimate? Is it surprising that the coefficient's standard error is larger?

Answer: The IV estimate is slightly larger. The standard error (0.013) is larger than the OLS standard error (0.006) but this is to expected from IV estimates.

- (g) v are the residuals from the model estimated in column 2. Test the null hypothesis that ltotexpend is exogenous.

Answer: When we obtain the reduced form residuals, say v , from column 2 and add them to the OLS regression in column 1, the robust t statistic on v is only 1.14. Therefore, we cannot reject the null hypothesis that ltotexpend is exogenous.

	OLS <i>sfood</i>	OLS log(<i>totexpend</i>)	IV <i>sfood</i>	OLS <i>sfood</i>
log(<i>totexpend</i>)	-0.146** (0.006)		-0.160** (0.013)	-0.160** (0.013)
<i>age</i>	0.002** (0.000)	0.005** (0.001)	0.002** (0.000)	0.002** (0.000)
<i>kids</i>	0.034** (0.005)	0.064** (0.018)	0.035** (0.005)	0.035** (0.005)
log(<i>income</i>)		0.478** (0.024)		
<i>v</i>				0.018 (0.015)
Constant	0.896** (0.027)	1.922** (0.114)	0.952** (0.054)	0.952** (0.054)
Observations	1519	1519	1519	1519
R^2	0.286	0.256	0.284	0.287

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

4. **(10 points)** Suppose you were interested in analyzing the impact of big-box retailers such as Walmart on the welfare of smaller mom and pops stores. Supposing your only datasource was a series of phonebooks for different counties and different years. Explain how you might test whether mom and pops stores are more likely to exit if a Walmart store opens up in the same county. Provide a possible regression equation and discuss any problems you will encounter (perhaps with potential solutions?) in carrying out this research project.

(Brief) Answer: A possible panel data regression equation might be:

$$exit_{ijt} = \beta_0 + \beta_1 Walmart_{jt} + u$$

where $exit_{ijt}$ equals 1 if mom and pop store i in county j exits in year t and $Walmart_{jt}$ is a dummy variable equal to 1 if a Walmart exists in county j in year t . If the existence of a Walmart increases the likelihood of exit then β_1 should be positive. A LPM, logit, or probit model can be used. One issue is that Walmart might locate in attractive counties that are thriving or will thrive, which might bias the estimate of β_1 upwards. A suitable instrument would therefore be needed. In class we discussed the “distance to Bentonville, Arkansas”

(Walmart's headquarters) as a possible instrument. This is likely to be correlated with the presence of a Walmart but uncorrelated with other characteristics in u .

Formulae

- For simple regression model: $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$
- Estimated slope parameter when regression equation passes through origin: $\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$
- $SSR = \sum_{i=1}^n \hat{u}_i^2$
- $SSE = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$
- $SST_j = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$
- $R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$
- For simple regression model: $se(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$
- $\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1-R_j^2)}$
- t statistic: $t_{\hat{\beta}_j} \equiv \frac{\hat{\beta}_j - \alpha_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} = t_{df}$
- 95% confidence interval: $P\left(\hat{\beta}_j - c_{0.05} \cdot se(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + c_{0.05} \cdot se(\hat{\beta}_j)\right) = 0.95$
- F statistic = $\frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$
- R^2 form of F statistic = $\frac{(R_{ur}^2 - R_r^2)/q}{(1-R_{ur}^2)/(n-k-1)}$
- $\bar{R}^2 = 1 - \frac{SSR/(n-k-1)}{SST/(n-1)}$
- Chow test statistic: $F = \frac{[SSR_P - (SSR_1 + SSR_2)]/(k+1)}{(SSR_1 + SSR_2)/[n-2(k+1)]}$

Cumulative Areas under the Standard Normal Distribution

<i>z</i>	0	1	2	3	4	5	6	7	8	9
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148

(continued)

<i>z</i>	0	1	2	3	4	5	6	7	8	9
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Examples: If $Z \sim \text{Normal}(0,1)$, then $P(Z \leq -1.32) = .0934$ and $P(Z \leq 1.84) = .9671$.

Source: This table was generated using the Stata® function `normprob`.

Critical Values of the *t* Distribution

Significance Level						
1-Tailed: 2-Tailed:		.10 .20	.05 .10	.025 .05	.01 .02	.005 .01
D e g r e e s o f F r e e d o m	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	30	1.310	1.697	2.042	2.457	2.750
	40	1.303	1.684	2.021	2.423	2.704
	60	1.296	1.671	2.000	2.390	2.660
	90	1.291	1.662	1.987	2.368	2.632
	120	1.289	1.658	1.980	2.358	2.617
	∞	1.282	1.645	1.960	2.326	2.576

Examples: The 1% critical value for a one-tailed test with 25 *df* is 2.485. The 5% critical value for a two-tailed test with large (> 120) *df* is 1.96.

Source: This table was generated using the Stata® function `invttail`.

F Values for $\alpha = 0.05$

d_2	d_1								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.3	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
inf	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38