Economics 113 Professor Spearot
Introduction to Econometrics
Fall 2008 – Final
Name Jane P. Answer Key

Final – 115 Points

You must answer all the questions. Please write your name on every page. The exam is closed book and closed notes. You may use calculators, but they must not be graphing calculators. Do not use your own scratch paper. And, have a nice holiday!!!!!!

You must show your work to receive full credit

Problem 1

a.) Congress is considering a "bailout" of the Detroit 3 Automakers. The worry is that if GM or Chrysler fail, then there would be severe consequences to the real economy. Although these numbers are fictitious, consider the following scenario. Suppose that the probability of a GM bankruptcy is 0.6 and the probability of a Chrysler bankruptcy 0.4. If these two events are independent of one another, what is the probability of GM OR Chrysler going bankrupt? **(10 Points)**

+5

$Pr(G \cup C) = Pr(G) + Pr(C) - Pr(G \cap C)$

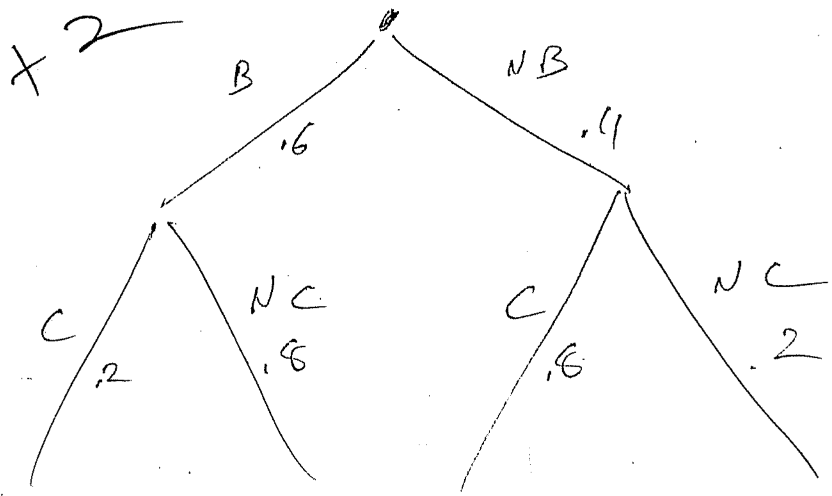
$= 0.6 + 0.4 - 0.6 \cdot 0.4$

$Pr(G \text{ or } C) = 0.76$

+5

b.) Now consider the effect of a bailout on the probability of economic collapse (again, these numbers are made-up). Suppose that the probability of a congressional bailout is 0.6. If congress approves a bailout, the probability of an economic collapse is 0.2. In contrast, if congress does not approve a bailout, the probability of an economic collapse is 0.8. Given that a collapse occurred, what is probability that congress approved a bailout? **(10 Points)**

+2



$$Pr(B|C) = \frac{Pr(B \cap C)}{Pr(C)}$$

+2

Using $Pr(C|B) = \frac{Pr(B \cap C)}{Pr(B)}$

+2

$\Rightarrow Pr(B \cap C) = Pr(C|B) \cdot Pr(B)$

$= 0.2 \cdot 0.6 = 0.12$

$Pr(C) = Pr(B \cap C) + Pr(NB \cap C)$ +2

$\quad .4 \cdot .8$

$= 0.12 + 0.32$

$Pr(C) = 0.44$

+2

$$Pr(B|C) = \frac{0.12}{0.44} = 0.273$$

2. For this question, we will examine the determinants of fuel efficiency. To do so, we have downloaded the Fuel Efficiency Database from the EPA (excluding niche performance brands, such as Ferrari) for the year 2008. Included in this database is every vehicle available for purchase in the United States, the measured fuel efficiency, and other characteristics.

To start the analysis, we run the following regression

$$MPG = \beta_0 + \beta_{displ} displ + \beta_{cyl} cyl + \beta_{auto} auto + u$$

Here, *MPG* is represents average miles per gallon travelled. Higher values imply greater efficiency. Further, *cyl* represents the number of cylinders in the engine, and *displ* represents the displacement of the engine in liters. Finally, *auto* is a dummy variable taking on the value of 1 if the vehicle has an automatic transmission and 0 if it has a standard transmission (stick shift).

Source	SS	df	MS
Model	15366.7919	3	5122.26398
Residual	7553.12472	1196	6.31532167
Total	22919.9167	1199	19.1158604

Number of obs	= 1200
F(3, 1196)	= 811.09
Prob > F	= 0.0000
R-squared	= 0.6705
Adj R-squared	= 0.6696
Root MSE	= 2.513

cmb	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
displ	-2.09063	.1343116	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx		
cyl	-.6174908	.0990547	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx		
auto	-.0072103	.1794679	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx		
_cons	30.44108	.2883117	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx		

a.) Using the 95% confidence level, test whether the coefficient on automatic, β_{auto}, is significantly different from zero. Please state your null and alternative hypotheses, and briefly interpret the result. **(10 Points)**

+1 $H_0: B_{auto} = 0$ $t_{stat} = \frac{-0.00721 - 0}{0.1795}$ +2

+1 $H_A: B_{auto} \neq 0$ $= -0.0402$

$t_{crit} = 1.96$ $|t_{stat}| < t_{crit}$ +2

Fail to reject H_0! +2

Having an automatic transmission has no significant effect of fuel efficiency +2 b.) Professor Spearot just purchased a car with 6 cylinders rather than 4. Holding all other attributes equal, what is the effect of purchasing a car with two extra cylinders on fuel efficiency? **(5 Points)**

$$\Delta MPG = \hat{\beta}_{cyl} \Delta Cyl \qquad +2$$
$$= -0.6175 \cdot 2$$
$$\boxed{\Delta MPG = -1.235} \qquad +3$$

c.) Also with 6 cylinders, Professor Spearot's car has an automatic transmission and an engine with 3.4 liters of displacement. Please **DERIVE** an estimating equation that will enable you to predict fuel efficiency for this car, and the standard error of the prediction. Please also write the commands required to generate any new variables in STATA. **(10 Points)**

$$\theta = \beta_0 + \beta_{disp} Displ + \beta_{cyl} Cyl + \beta_{auto} Auto$$
$$\theta = \beta_0 + \beta_{disp} \cdot 3.4 + \beta_{cyl} \cdot 6 + \beta_{auto} \cdot 1 \qquad +4$$
$$\beta_0 = \theta - \beta_{disp} \cdot 3.4 - \beta_{cyl} \cdot 6 - \beta_{auto} \cdot 1$$

$$MPG = \theta - \beta_{disp} \cdot 3.4 - \beta_{cyl} \cdot 6 - \beta_{auto} \cdot 1 + \beta_{disp} \cdot disp + \beta_{cyl} Cyl + \beta_{auto} Auto + u$$

$$MPG = \theta + \beta_{disp}(Displ - 3.4) + \beta_{cyl}(Cyl - 6) + \beta_{auto}(Auto - 1) + u \qquad +4$$

```
gen displ_3.4 = displ - 3.4
gen cyl_6 = Cyl - 6
gen Auto_1 = Auto - 1
```
+2

The model in part (a) is much too basic. To remedy this, I add in *truck* and *domestic*, where the former is a dummy variable equal to 1 if the vehicle is a truck (0 otherwise), and the latter is a dummy variable equal to 1 if the vehicle is domestic (0 otherwise).

$$MPG = \beta_0 + \beta_{displ} displ + \beta_{cyl} cyl + \beta_{auto} auto + \beta_{truck} truck + \beta_{dom} domestic + u$$

Note that the classification "truck" includes SUVs and vans. The results from estimating this equation are below:

```
      Source |       SS       df       MS              Number of obs =    1200
-------------+------------------------------           F(  5,  1194) =  612.55
       Model |  16490.9879     5  3298.19758           Prob > F      =  0.0000
    Residual |  6428.92875  1194  5.38436243           R-squared     =  0.7195
-------------+------------------------------           Adj R-squared =  0.7183
       Total |  22919.9167  1199  19.1158604           Root MSE      =  2.3204

------------------------------------------------------------------------------
         cmb |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       displ |  -1.379712   .1386449   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
         cyl |  -.9957897   .0974275   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
        auto |   .3522417   .1682961   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       truck |  -2.230827   .1563322   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
    domestic |   .1960408   .1522806   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       _cons |    30.8768   .2699396   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
------------------------------------------------------------------------------
```

d.) Is this model preferred to the model in (a)? If a hypothesis test is warranted, test this hypothesis at the 95% level, stating your null and alternative hypotheses. If not, provide other evidence for your conclusion. **(10 Points)**

$H_0: B_{truck} = B_{dom} = 0$ +1

$F_{crit} = 3.00$ +1

H_A: H_0 not true +1

+1

$$F_{stat} = \frac{\frac{SSR_R - SSR_{UR}}{q}}{\frac{SSR_{UR}}{dof}} = \frac{\frac{7553.1 - 6428.9}{2}}{\frac{6428.9}{1194}} = 104.39$$

$104.39 > 3$ +1

Reject H_0 in favor of H_A +2

e.) Using a 90% confidence level, produce a confidence interval for the coefficient on *truck*, β_{truck}. Briefly interpret this interval. **(10 Points)**

$CI_{Truck} = \hat{\beta}_{truck} \pm se(\hat{\beta}_{truck}) \cdot t_{crit}(90)$

$\hat{\beta}_{truck} = -2.231$, $se = 0.156$, $t_{crit} = 1.64$ +1

$-2.49 < \beta_{truck} < -1.98$ +1

With 90% confidence, holding all else equal, trucks decrease fuel economy between 2.49 and 1.98 miles per gallon +2

f.) Trucks tend to be very heavy, and heavy vehicles tend to be less fuel-efficient. Since the EPA database does not include vehicle weight as a variable, weight is an omitted variable. In what direction is the bias? Can I be confident that the coefficient on *truck*, β_{truck}, will still be negative after accounting for this bias? **(10 Points)**

$MPG = \ldots \beta_{truck} Truck + u$ +3

Negative (downward) Bias +2

Since this bias allows for the possibility that $\beta_{truck} > 0$, we cannot rule it out. Thus, β_{truck} is not necessarily negative +5 g.) Domestic vehicles are often claimed to be less fuel-efficient than their equivalent foreign counterparts. Interpret the coefficient on *domestic*, β_{dom}, and test whether it is significantly less than zero (you do not need a confidence level for this question). Is there evidence for such a claim? **(5 Points)**

None For this, the null is $B_{dom} \geq 0$ and the alternative is $B_{dom} < 0$. Since $\hat{B}_{dom} > 0$, then we fail to reject the null.

+5

h.) Using the regression from (d), Professor Spearot claims that the fuel efficiency for cars with automatic transmissions is different from those with standard transmissions. What is the probability that he is wrong? **(10 Points)**

$$t_{stat} = \frac{0.352 - 0}{0.168} = 2.09$$

+2

$$P_{value} = Pr(|T| > 2.09)$$
$$= 2(1 - Pr(T < 2.09))$$
$$= 2(1 - 0.9817)$$
$$= 0.0366$$

There is a 3.66% ~~chance~~ chance of being wrong

Engines are an important determinant of fuel-efficiency, and thus, there may be an important interaction between engine displacement *(displ)* and the number of cylinders *(cyl)* contained within the engine. To examine this possibility, we estimate the following equation:

$$MPG = \beta_0 + \beta_{displ} displ + \beta_{cyl} cyl + \beta_{dc} displ * cyl + \beta_{auto} auto + \beta_{truck} truck + \beta_{dom} domestic + u$$

The results from estimating this equation are below:

```
      Source |       SS       df       MS              Number of obs =    1200
-------------+------------------------------           F(  6,  1193) =  581.73
       Model |  17081.4864     6  2846.9144            Prob > F      =  0.0000
    Residual |  5838.43025  1193  4.89390633           R-squared     =  0.7453
-------------+------------------------------           Adj R-squared =  0.7440
       Total |  22919.9167  1199  19.1158604           Root MSE      =  2.2122

------------------------------------------------------------------------------
         cmb |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       displ |  -3.665261   .2465046   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
         cyl |  -2.243822   .1467528   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   displ*cyl |    .341223    .031064   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
        auto |   .4683869   .1607962   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       truck |  -1.963128   .1510215   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
    domestic |    .187432   .1451816   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       _cons |   38.37429    .729455   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

i.) **Derive** the effect of engine displacement (*displ*) on fuel efficiency (*MPG*). Is the interaction between engine displacement and cylinders significant? Test this hypothesis at the 99% level. **(10 Points)**

$$\frac{\partial MPG}{displ} = \beta_{displ} + \beta_{dc} cyl$$

+4

$H_0: \beta_{dc} = 0$

$H_A: \beta_{dc} \neq 0$

$$t_{stat} = \frac{0.3412 - 0}{0.0311} = 10.97$$

+4

$t_{crit} = 2.57$

$|t_{stat}| > t_{crit}$

There is a significant interaction!!

+2

j.) Professor Spearot's car is black, but color does not affect fuel efficiency. Suppose that he includes *Color* in any of the above regressions. What (if any) are the effects of adding *color*? **(5 Points)**

Standard errors (variance) will be higher.
or Estimates less precise +5

Partial DOF ↓ +2

k.) Suppose that the variance of unobservable factors affecting fuel efficiency increases with engine displacement. What problem is this and how do we correct for it? **(5 Points)**

Heteroskedasticity +3
Robust +2

l.) Suppose I run the following regression predicting the probability that a specific vehicle is domestic:

$$Domestic = \beta_0 + \beta_{displ} displ + \beta_{auto} auto + \beta_{truck} truck + u$$

I notice in the results that some predictions are negative. What should be done to correct for this problem? **(5 Points)**

Use Probit/Logit +5

Extra Credit (5 points)

Suppose that instead of the regression in (i), I ran the following:

$$MPG = \beta_0 + \beta_{displ} displ + \beta_{dc} displ * cyl + \beta_{auto} auto + \beta_{truck} truck + \beta_{dom} domestic + u$$

What problem am I introducing and why?

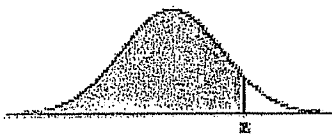
or E(u|X)

Omitted Variable Bias +2

cyl is in u +3

correlated with MPG

" " displ · cyl



Normal Distribution from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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

TABLE G.3b

5% Critical Values of the *F* Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom: 1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator $df = 4$ and large denominator $df\ (\infty)$ is 2.37.
Source: This table was generated using the Stata® function invFtail.

TABLE G.3b

5% Critical Values of the F Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom: 1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator $df = 4$ and large denominator df (∞) is 2.37.
Source: This table was generated using the Stata® function invFtail.