Fundamentals of Statistical Modeling (VT20)

Andrea Discacciati
Karolinska Institutet
Stockholm, Sweden

Lab 2 (Extra material on convenient parametrizations)

Load the dataset and the `mlci` command

```
. version 14
. use https://raw.githubusercontent.com/anddis/fsm/master/data/lab2.dta, clear
. run https://raw.githubusercontent.com/anddis/fsm/master/do/mlci.do
```

Exercise 1

So far, we've used the gamma distribution parametrized by parameters α and β. They are not interpretable.

```
. local alpha = "exp({theta1})"
. local beta = "exp({theta2})"
. local f = "gammaden(`alpha´, `beta´, 0, y)"
. mlexp (ln(`f´))
initial:       log likelihood =     -<inf>  (could not be evaluated)
feasible:      log likelihood = -96414.257
rescale:       log likelihood = -13891.173
rescale eq:    log likelihood = -13891.173
Iteration 0:   log likelihood = -13891.173
Iteration 1:   log likelihood = -8165.6417
Iteration 2:   log likelihood = -8160.8897
Iteration 3:   log likelihood = -8160.8781
Iteration 4:   log likelihood = -8160.8781
Maximum likelihood estimation
Log likelihood = -8160.8781                     Number of obs     =      1,432
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/theta1	.3906123	.033984	11.49	0.000	.3240049	.4572196
/theta2	4.349872	.0403414	107.83	0.000	4.270804	4.42894

```
. mlci exp /theta1
 1.477885   95% CI: 1.382654, 1.579676
. mlci exp /theta2
 77.46854   95% CI: 71.57918, 83.84246
```

The mean of a gamma distribution is equal to $\alpha\beta$ (see Wikipedia).

```
. di exp(_b[/theta1])*exp(_b[/theta2])
114.48962
```

Sometimes it can be useful to parametrize the gamma distribution in such a way that one of its 2 parameters is equal to (a transform of) the mean.

We define a parameter for the mean: $E[Y] \equiv \eta = \alpha\beta$. Then, $\alpha = \eta/\beta$ (see the slide "Convenient Parametrizations" for an ananlogous example with the log-normal distribution).

Note that the log-likelihood of this model is identical to the one of the previous one. No surprise: the model is exactly the same, it's just its parametrization that changed.

```
. local eta = "exp({theta1})" // We constrain the mean to be strictly positive

. local beta = "exp({theta2})"

. local alpha = "`eta´ / `beta´"

. local f = "gammaden(`alpha´, `beta´, 0, y)"

. mlexp (ln(`f´))
initial:       log likelihood =     -<inf>  (could not be evaluated)
feasible:      log likelihood = -100156.17
rescale:       log likelihood = -8730.8329
rescale eq:    log likelihood = -8730.8329
Iteration 0:   log likelihood = -8730.8329
Iteration 1:   log likelihood = -8309.9438
Iteration 2:   log likelihood = -8161.7729
Iteration 3:   log likelihood = -8160.8785
Iteration 4:   log likelihood = -8160.8781

Maximum likelihood estimation

Log likelihood = -8160.8781                     Number of obs     =      1,432

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     /theta1 |   4.740484   .0217374   218.08   0.000     4.697879    4.783088
     /theta2 |   4.349871   .0403413   107.83   0.000     4.270803    4.428938
------------------------------------------------------------------------------

. mlci exp /theta1 // This gives me the MLE of eta (the mean) with 95% CI
 114.4896   95% CI: 109.7142, 119.4728

. mlci exp /theta2
 77.46844   95% CI:  71.5791, 83.84235
```