University of Essex
2015-16

Department of Government
TA: Stavros Poupakis

GV903: Advanced Research Methods

Class 14 Estimation

Simulate `x` drawn from a Normal distribution of mean 0 and and variance 4, with sample size 500.

1. Using the method of moments, estimate the population mean.
2. Create a function as below

```
loglik <- function(a, mu, sigma2){
  sum( log( dnorm(a, mu, sqrt(sigma2)) ) )
}
```

   this is essentially a loglikelihood function of the Normal distribution. Let `a` to be your `x`, `sigma2` the estimate of the population variance and try several different values for `mu`.

3. Write a function that will calculate the loglikelihood again but write explicitly the pdf, that is without using the `dnorm` function. Check that you get the same results as before.
4. For the two vectors below, for every element of `play` store in the respective element of `ll` the value of the loglikelihood. Plot these two vectors.

```
play <- seq(-10,10,by=0.1)
ll <- rep(NA,length(play))
```

   Can we look closer?

5. Create a function that calculates the likelihood, e.g. instead of `sum` take `prod` and don't use any logs. Why don't we just use the likelihood and we prefer to take logs?
6. Create the function

```
mle <- function(par, a){
   mu <- par[1]
   s2 <- exp(par[2])

   ll <- loglik(a, mu, s2)
  -ll
}
```

   and use `optim(c(0,1), mle, a=x, method="BFGS")`. We will discuss more on this in class.

7. For a sample size of 1,000 simulate $x_i \sim N(0, 10^2)$, $\epsilon_i \sim N(0, 9^2)$ and create $y_i = 2 + 3x_i + \epsilon_i$. Use the `lm` function to get the OLS estimates.
8. Create 3 functions, the one will return the OLS using the well-known formula which you will write explicitly. For the other two, one will return the residual sum of squares and the other the loglikelihood. One needs to be minimised and one to be maximized (using `optim`) to get the estimates. For the MLE use also the option `hessian=T`.