Tutorial: Normal distribution in Stata

Using Stata to calculate Normal probabilities

Suppose Z is a standard normal random variable. When $Z \sim$ Normal(0, 1),

`normal(z)`	returns the cumulative standard normal distribution
`normalden(z)`	returns the standard normal density

Example: Ozone Designation Following the Clean Air Act Amendments of 1997

From 2001-2003, the Environmental Protection Agency (EPA) monitored ozone levels at monitors across the United States. One criteria for ozone was that the ozone levels (defined as the average fourth highest daily maximum ozone over the three year period) could not exceed 80ppb. Regulatory actions were taken if the ozone levels exceeded this threshold.

Among monitors in the Southeast, the average ozone level was 45.2 ppb, with standard deviation 6.3 ppb. Ozone levels are usually modeled using the normal distribution. We assume that this distribution is reasonable in our application.

Define X as ozone level at a monitor. $X \sim N(45.2, 39.7)$, or, equivalently, $X \sim N(45.2, 6.3^2)$.

1. What is the expected ozone level at a randomly sampled monitor?

   45.2 ppb

2. What is the typical departure ozone levels from this mean number?

   6.3 ppb

3. Why do you think Stata named the normal density function `normalden`, rather than `normalp`, which would seemingly be more consistent with the binomial and Poisson commands?

   The normal distribution is continuous, and therefore `normalden` does not return a probability, but rather a density function.

4. Why do you think Stata only calculates probabilities with respect to the standard normal, or N(0,1), distribution?

   I don't know the answer to this. Seems pretty inconvenient.

5. What is the probability that a randomly selected monitor has ozone levels exceeding 80 ppb?

   First, standardize:

   $P(\frac{X-45.2}{6.3} > \frac{80-45.2}{6.3}) = P(Z > 5.524)$

```
. di 1 - normal(5.524)
1.657e-08
```

6. Provide an interpretation of the following command:

```
. di normalden(0)
.39894228
```

0.399 is the value of the normal density function at 0. It has no interpretation in terms of probability.