Statistics and Data Analysis

Lecture 4

Lecture by Dr. Zohar Yakhini
Typeset by Steven Karas

2016-11-29
Last edited 17:58:35 2016-12-28

1 Normal Distribution

Sometimes called the Gaussian distribution. Colloquially referred to as a bell curve.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

1.1 Standard Normal

A special case of the Normal distribution with $\mu = 0$ and $\sigma = 1$.

1.2 CDF

The function is not integrable. Lookup tables and numeric computation are used instead.

Example: Bicycle Oil In a bicycle shop, they sell oil. When the stock of oil drops below 20, they order new stock. However, the delivery takes a full day to arrive. What are the chances of running out of stock while waiting for the delivery?

Let the daily demand be $\sim N(\mu = 15, \sigma = 6)$ bottles of oil.

1.3 Z-Shift

$$Z = \frac{x - \mu}{\sigma}$$

$$\frac{dz}{dx} = \frac{1}{\sigma}$$

$$\frac{1}{2\sqrt{\pi}} \int_I e^{\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{1}{2\sqrt{\pi}} \int_I e^{\frac{z^2}{2}} dz$$

Using this technique, we can transform any integral on any normal distribution to the standard normal.

Example: Bicycle Oil

$$z = \frac{(x-\mu)}{\sigma} = \frac{20-15}{6} = 0.83$$

From our reference tables/tools, we get that this is 0.2033

1.4 Sampling from a Uniform distribution to any distribution

Let f be a CDF. $f(x) = y$, and because f is monotonous, $f^{-1}(y) = x$.

Therefore, $P(f^{-1}(\text{RAND}) < x) = P(\text{RAND} < y) = y = f(x)$

In short, to generate, we draw uniform random numbers and use the inverse CDF to generate the values according to the distribution.

1.5