CmpE 343 Lecture Notes
6: Continuous Distributions

Ethem Alpaydın

October 28, 2014

1 Introduction

Frequently, certain distributions arise in different applications, and they are given names and parameterized. Now we discuss the distributions where the random variable is continuous.

2 Uniform Distribution

If the random variable X is uniform distributed in the interval $[L, U]$, we write the probability density as

$$f(x; L, U) = \begin{cases} \frac{1}{U-L} & \text{if } L < x < U, \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

This is because, to be a legitimate probability density, we need to have $\int f(x)dx = 1$.

The expected value and variance of the uniform distributed X are given as

$$\begin{aligned} E[X] &= \frac{L+U}{2} \\ \text{Var}(X) &= \frac{(U-L)^2}{12} \end{aligned} \tag{2}$$

3 Normal (Gaussian) Distribution

This is the bell-shaped distribution used frequently in various types of applications. Its probability density function is given as

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty \tag{3}$$

The center of the bell is given by the mean, μ—it is also the *mode* of the distribution, and the variance, σ^2 defines how fast the curve descends symmetrically on both sides as we move away from the mean.

If $X \sim N(\mu, \sigma^2)$ then

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1) \tag{4}$$

This is called *z transformation* and $N(0, 1)$ is called the *unit normal distribution*. Most of the density of a normal distribution lies close to its mean, for example, $P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2) = 0.95$ (Chebyshev's bound is 0.75).

Linear combination of independent normalliy distributed random variables is also normally distributed. If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ and the two are independent then

$$c_1X_1 + c_2X_2 \sim N(c_1\mu_1 + c_2\mu_2, c_1\sigma_1^2 + c_2\sigma_2^2)$$

This can be generalized to more than two such variables.

4 Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be a set of independent and identicall distributed (iid) random variables each with mean μ and variance σ^2, then their sum is approximately normal distributed with mean $n\mu$ and variance $n\sigma^2$. Or

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \dot{\sim} N(0,1) \tag{5}$$

The power of the central limit theorem lies in the fact that X_i can be from *any* distribution. As long as X_i are all independently drawn from the same distribution, their sum (or average) is approximately normal. For example, we know that the Binomial can be written as the sum of independent Bernoullis, so from the central limit theorem we can see that $X \sim \text{Binomial}(n, p_0)$ is approximately normal with mean np_0 and variance $np_0(1-p_0)$, or

$$\frac{X - np_0}{\sqrt{np_0(1-p_0)}} \dot{\sim} N(0,1)$$

5 Chi-Squared Distribution

If $Z \sim N(0,1)$ then Z^2 is Chi-Squared with one degree of freedom, denoted as χ_1^2. The sum of independent Chi-Squared distributed random variables is also Chi-Squared with its degree of freedom equal to the sum of the degrees of freedom.

For example, if $X_i \sim N(\mu, \sigma^2), i = 1, \ldots, N$, then

$$\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_1^2$$

and for iid X_i

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_n^2$$