Chris Piech
CS109

Handout #16
April 18th, 2016

Gaussian

Normal Random Variable

The single most important random variable type is the Normal (aka Gaussian) random variable, parametrized by a mean (μ) and variance (σ^2). If X is a normal variable we write $X \sim \mathcal{N}(\mu, \sigma^2)$. The normal is important for many reasons: it is generated from the summation of independent random variables and as a result it occurs often in nature. Many things in the world are not distributed normally but data scientists and computer scientists model them as Normal distributions anyways. Why? Because it is the most entropic (conservative) distribution that we can apply to data with a measured mean and variance.

Properties

The Probability Density Function (PDF) for a Normal is:

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

By definition a Normal has $E[X] = \mu$ and $Var(X) = \sigma^2$.

If X is a Normal such that $X \sim \mathcal{N}(\mu, \sigma^2)$ and Y is a linear transform of X such that $Y = aX + b$ then Y is also a Normal where $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

There is no closed form for the integral of the Normal PDF, however since a linear transform of a Normal produces another Normal we can always map our distribution to the “Standard Normal” (mean 0 and variance 1) which has a precomputed Cumulative Distribution Function (CDF). The CDF of an arbitrary normal is:

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Where Φ is a precomputed function that represents that CDF of the Standard Normal.

Projection to Standard Normal

For any Normal X we can define a random variable $Z \sim \mathcal{N}(0, 1)$ to be a linear transform

$$\begin{aligned} Z &= \frac{X-\mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \\ &\sim \mathcal{N}\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{\sigma^2}{\sigma^2}\right) \\ &\sim \mathcal{N}(0, 1) \end{aligned}$$

Using this transform we can express $F_X(x)$, the CDF of X, in terms of the known CDF of Z, $F_Z(x)$. Since the CDF of Z is so common it gets its own Greek symbol: $\Phi(x)$

$$\begin{aligned} F_X(x) &= P(X \le x) \\ &= P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) \\ &= P\left(Z \le \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

The values of $\Phi(x)$ can be looked up in a table. We also have an online calculator.

Example 1

Let $X \sim \mathcal{N}(3, 16)$, what is $P(X > 0)$?

$$P(X > 0) = P\left(\frac{X-3}{4} > \frac{0-3}{4}\right) = P\left(Z > -\frac{3}{4}\right) = 1 - P\left(Z \le -\frac{3}{4}\right)$$
$$= 1 - \Phi(-\frac{3}{4}) = 1 - (1 - \Phi(\frac{3}{4})) = \Phi(\frac{3}{4}) = 0.7734$$

What is $P(2 < X < 5)$?

$$P(2 < X < 5) = P\left(\frac{2-3}{4} < \frac{X-3}{4} < \frac{5-3}{4}\right) = P\left(-\frac{1}{4} < Z < \frac{2}{4}\right)$$
$$= \Phi(\frac{2}{4}) - \Phi(-\frac{1}{4}) = \Phi(\frac{1}{2}) - (1 - \Phi(\frac{1}{4})) = 0.2902$$

Example 2

You send voltage of 2 or -2 on a wire to denote 1 or 0. Let X = voltage sent and let R = voltage received. $R = X + Y$, where $Y \sim \mathcal{N}(0, 1)$ is noise. When decoding, if $R \ge 0.5$ we interpret the voltage as 1, else 0. What is $P(\text{error after decoding}|\text{original bit} = 1)$?

$$P(X + Y < 0.5) == P(2 + Y < 0.5) = P(Y < -1.5) = \Phi(-1.5) = 1 - \Phi(1.5) \approx 0.0668$$

Binomial Approximation

You can use a Normal distribution to approximate a Binomial $X \sim Bin(n, p)$. To do so define a normal $Y \sim (E[X], Var(X))$. Using the Binomial formulas for expectation and variance, $Y \sim (np, np(1-p))$. This approximation holds for large n. Since a Normal is continuous and Binomial is discrete we have to use a continuity correction to discretize the Normal.

$$P(X = k) \sim P\left(k - \frac{1}{2} < Y < k + \frac{1}{2}\right) = \Phi\left(\frac{k - np + 0.5}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - np - 0.5}{\sqrt{np(1-p)}}\right)$$

Example 3

100 visitors to your website are given a new design. Let X = # of people who were given the new design and spend more time on your website. Your CEO will endorse the new design if $X \ge 65$. What is $P(\text{CEO endorses change}|\text{it has no effect})$?

$E[X] = np = 50$. $Var(X) = np(1-p) = 25$. $\sigma = \sqrt{(Var(X))} = 5$. We can thus use a Normal approximation: $Y \sim \mathcal{N}(50, 25)$.

$$P(X \ge 65) \approx P(Y > 64.5) = P\left(\frac{Y - 50}{5} > \frac{64.5 - 50}{5}\right) = 1 - \Phi(2.9) = 0.0019$$

Example 4

Stanford accepts 2480 students and each student has a 68% chance of attending. Let X = # students who will attend. $X \sim Bin(2480, 0.68)$. What is $P(X > 1745)$?

$E[X] = np = 1686.4$. $Var(X) = np(1-p) = 539.7$. $\sigma = \sqrt{(Var(X))} = 23.23$. We can thus use a Normal approximation: $Y \sim \mathcal{N}(1686.4, 539.7)$.

$$P(X > 1745) \approx P(Y > 1745.5) = P\left(\frac{Y - 1686.4}{23.23} > \frac{1745.5 - 1686.4}{23.23}\right) = 1 - \Phi(2.54) = 0.0055$$