Probability Theory for EOR

Some special continuous random variables
III (Normal)
Part A

Some **continuous random variables** are associated very special/ubiquitous **distributions (PDFs)**, they get their own names!

Definition (PDF of continuous real-valued r.v.)

The **probability density function (PDF)** of a **continuous** real-valued r.v. is a non-negative function f_X on the real line such that via the Riemann integral:

$$\int_{-\infty}^{x} f_X(s)ds = P(X \leq x).$$

For a **continuous** random variable with differentiable CDF F_X, conventionally, $f_X(x) = F_X'(x)$.

Normal/Gaussian!! Part A.

Normal distribution

- Suppose there is a super server receiving many requests from independent computers $1, \cdots, n$, the number of requests from each computer within a time interval of unit length, $N_i, i = 1, \cdots, n$, would follow i.i.d. Pois(1).
- One is wondering about the distribution of the number of the total requests when n is getting larger and larger, and decided to look at the following random varaible (a rescaled of the total requests shifited by the average):

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (N_i - EN_i),$$

which is $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (N_i - 1)$.

- Note that $\sum_{i=1}^{n} N_i$ is a sum of i.i.d. Pois(1)-distributed r.v.'s, which is still a Poisson distribution (Pois(n)), and thus we can calculate the CDF function of the term $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (N_i - 1)$:

$$P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (N_i - 1) \leq m\right) = P(\sum_{i=1}^{n} N_i \leq n + \sqrt{n}m) = \sum_{i=0}^{\lfloor n+\sqrt{n}m \rfloor} e^{-n} \frac{n^i}{i!},$$

which surprisingly has a limit as $n \to \infty$:

$$\int_{-\infty}^{m} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds.$$

- **The limit is true not only for $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (N_i - EN_i)$ but also valid for more general choices of X_i's (zero mean and unit varaince): $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - EX_i)$.**

Normal distribution: $N(\mu, \sigma^2)$

- A continuous r.v. Z is said to have the *standard normal distribution* $N(0,1)$ (mean zero and varaince one) if its PDF ψ is given by

$$\psi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

with corresponding CDF $\Psi(z) = \int_{-\infty}^{z} \psi(z)dz$.

- As for $X = \mu + \sigma Z$ ($\sigma > 0$), it is said to have the *normal distribution* $N(\mu, \sigma^2)$ with mean μ and variance σ^2.

  The PDF of X:

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

- $F_X(x) = P(\mu + \sigma Z \leq x) = \Psi(\frac{x-\mu}{\sigma}), f_X(x) = F'_X(x) = \psi(\frac{x-\mu}{\sigma})\frac{1}{\sigma}$.

$\psi(z)$ is a proper PDF

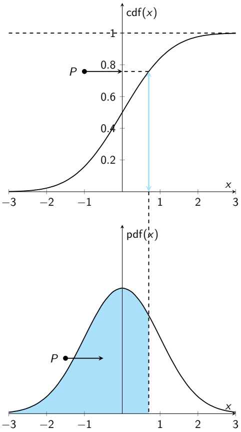
- Non-negative.
- Integrate to one:

$$\left(\int_{-\infty}^{\infty}\psi(z)dz\right)^2 = \left(\int_{-\infty}^{\infty}\psi(x)dx\right)\left(\int_{-\infty}^{\infty}\psi(y)dy\right)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi(x)\psi(y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy$$

$$= \int_{0}^{2\pi}\int_{0}^{\infty}\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdrd\theta = \int_{0}^{2\pi}\frac{1}{2\pi}d\theta = 1$$

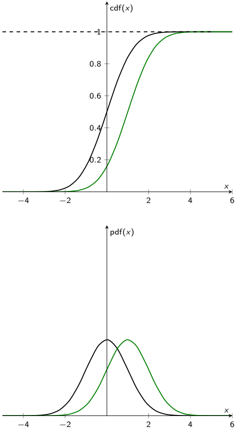
Standard Normal Distribution $N(0,1)$ CDF, PDF



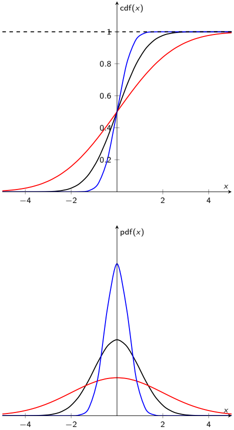
- Symmetry of PDF ($\psi(z) = \psi(-z)$); of CDF ($\Psi(z) = 1 - \Psi(-z)$); of Z and -Z.

Normal Distribution $N(\mu, \sigma^2)$ CDF, PDF

$\sigma = 1, \mu =$(black 0), (green 1)



$\mu = 0, \sigma =$(red 2), (black 1), (blue 1/2)



Expectation and Varaince

$Z \sim N(0,1)$, $X \sim N(\mu, \sigma^2)$.

- The expectation EZ.

$$\mathrm{E}[Z] = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx = 0.$$

- The varaince VZ.

$$\begin{aligned}
\mathrm{E}[Z^2] &= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx = 2 \int_{0}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \\
&= \frac{2}{\sqrt{2\pi}} \left(-xe^{-x^2/2}|_0^{\infty} + \int_0^{\infty} e^{-\frac{x^2}{2}} \, dx \right) \\
&= \frac{2}{\sqrt{2\pi}} \left(\sqrt{2\pi} \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \right) = \frac{2}{\sqrt{2\pi}} \left(\sqrt{2\pi} \frac{1}{2} \right) \\
&= 1. \\
\mathrm{V}[Z] &= \mathrm{E}[Z^2] - (\mathrm{E}[Z])^2 = 1.
\end{aligned}$$

- $EX = E(\mu + \sigma Z) = \mu$, $VX = V(\mu + \sigma Z) = \sigma^2$.