Preliminary – Central Limit Theorem and the Gaussian Distribution

Gaussian distribution

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

- $x \sim N(\mu, \sigma^2)$
- $\mathbb{E}[x] = \mu$ (the mean)
- $Var[x] = \sigma^2$ (the variance $\sigma^2 > 0$)

* Sum of independent gaussian random variables $x_1 \sim N(\mu_1, \sigma_1^2)$ and $x_2 \sim N(\mu_2, \sigma_2^2)$ are independent, then $x_1 + x_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

* Moment generating function and moments

- The moment generating function of $x \sim N(\mu, \sigma)$ is

$$M_x(t) = \mathbb{E}[\exp(tx)] = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

- The moments of $x \sim N(\mu, \sigma)$ is given by

$$\mathbb{E}[(x-\mu)^k] = \begin{cases} \sigma^k (k-1)!! & \text{, if } k \text{ is even} \\ 0 & \text{, if } k \text{ is odd} \end{cases}$$

where $(k-1)!! = (k-1)\times(k-3)\times\cdots\times 3\times 1$, the double factorial of $k-1$.

Gaussian vectors

$$x = \Sigma^{1/2} y + \mu$$

- $\mu \in \mathbb{R}^d$
- $\Sigma \in \mathbb{R}^{d\times d}$ is positive semi-definite matrix
- $x = [x_1, \cdots, x_d]^T \in \mathbb{R}^d$ and $x \sim N(\mu, \Sigma)$
- $y = [y_1, \cdots, y_d]^T \in \mathbb{R}^d$ and $y_1, \cdots, y_n \overset{iid}{\sim} N(0,1)$

* $\mathbb{E}[x] = \mu$

* $\mathbb{E}[(x-\mu)(x-\mu)^T] = \Sigma$

* If $\mu = 0$, $\Sigma = I_d$, x is a standard Gaussian random vector.

The joint probability distribution function x is given by

$$P_x(x) = \frac{1}{\sqrt{(2\pi)^d \cdot \det(\Sigma)}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

if Σ is positive semi-definite.

Linear transformation of a Gaussian

For every $\mu \in \mathbb{R}^d$, every PSD matrix $\Sigma \in \mathbb{R}^{d\times d}$ and every matrix $A \in \mathbb{R}^{d\times d}$. If $x \sim N(\mu, \Sigma)$, then $Ax \sim N(A\mu, A\Sigma A^T)$.

Rotational invariance of the standard Gaussian distribution

If $x \sim N(0, I_d)$ is standard Gaussian and $U \in \mathbb{R}^{d\times d}$ is an orthogonal matrix, then Ux is also standard Gaussian.

Orthogonal projections of a standard Gaussian

If $x \sim N(0, I_d)$ is standard Gaussian and $v_1, \cdots, v_d \in \mathbb{R}^d$ are mutually orthogonal vectors, then the random variables $v_1^T x, \cdots, v_d^T x$ are mutually independent.

Jointly Gaussian random variables

Fact 1: If $x = My$ for some matrix M and y is Gaussian, then x and y are jointly Gaussian.

Fact 2: Suppose x and y are independent random vectors such that both x and y are individually Gaussian, then x and y are jointly Gaussian.

Fact 3: Two jointly Gaussian vectors x, y are independent if and only if $\mathbb{E}[xy^T] = \mathbb{E}[x] \cdot \mathbb{E}[y]^T$. This condition is necessary for any two independent random vectors.

* Two jointly distributed random vectors x and y can be individually Gaussian, but not jointly Gaussian.

Example: $A \sim N(0,1)$

$$B = \begin{cases} A & \text{if } |A| < t \\ -A & \text{if } |A| \geq t \end{cases}$$

$$\mathbb{E}[AB] = \mathbb{P}(|A|<t) \cdot \mathbb{E}[A^2 \mid |A|<t] - \mathbb{P}(|A| \geq t) \cdot \mathbb{E}[A^2 \mid |A| \geq t]$$

$$\left.\begin{array}{l} - \; t>0, \ \mathbb{P}(A<t) \cdot \mathbb{E}[A^2 \mid |A|<t] = \frac{1}{2} \\ - \; \mathbb{E}[A^2] = 1, \ \mathbb{P}(|A| \geq t) \cdot \mathbb{E}[A^2 \mid |A| \geq t] = \frac{1}{2} \end{array}\right\}$$

~~A & B independent~~

contradict!

Central limit theorem

Let $(x_n)_{n \geq 1}$ be a sequence of independent and identically distributed (iid) random variables with mean μ and variance $0 < \sigma^2 < \infty$. For every $n \geq 1$, we define the sample average $\bar{x}$ as $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$.

* The law of large number implies that the sequence $(\bar{x}_n)_{n \geq 1}$ converges almost surely to the expected value μ.

* The CLT characterizes how the sample average $\bar{x}_n$ deviates from the expected value μ for large n.

More precisely, the CLT states that the sequence $\frac{\sqrt{n}(\bar{x}_n - \mu)}{\sigma}$ converges in distribution to a standard Gaussian random variable, i.e., for every $t \in \mathbb{R}$ we have

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\sqrt{n}(\bar{x}_n - \mu)}{\sigma} \leq t \right) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$

* The distribution of the aggregate of a large number of independent and identically distributed small random phenomena is always very close to a Gaussian.