Probability - Gaussian Distribution

Probability

	Discrete Random Variable	Continuous Random Variable
Probability	$P(X = x_i) = P(x_i)$	$P(a \le x \le b) = \int_a^b f(x)dx$
Expectation	$E(X) = \sum_{i=1}^{m} x_i P(x_i)$	$E(X) = \int_{-\infty}^{\infty} xf(x)dx$
Variance	$Var(X) = \sum_{i=1}^{m} (x_i - \mu)^2 P(x_i)$	$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$

where $f(x)$ is the probability density function.

Probability

$$Var(X)=\sum_{i=1}^{m}(x_i-\mu)^2P(x_i)$$
$$=\sum_{i=1}^{m}(x_i^2-2\mu x_i+\mu^2)P(x_i)$$
$$=\sum_{i=1}^{m}x_i^2P(x_i)-2\mu\sum_{i=1}^{m}x_iP(x_i)+\mu^2\sum_{i=1}^{m}P(x_i)$$
$$=\sum_{i=1}^{m}x_i^2P(x_i)-2\mu^2+\mu^2$$
$$=\sum_{i=1}^{m}x_i^2P(x_i)-\mu^2$$
$$=E(X^2)-[E(X)]^2$$

$$Var(X)=\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx$$
$$=\int_{-\infty}^{\infty}(x^2-2\mu x+\mu^2)f(x)dx$$
$$=\int_{-\infty}^{\infty}x^2f(x)dx-2\mu\int_{-\infty}^{\infty}xf(x)dx+\mu^2\int_{-\infty}^{\infty}f(x)dx$$
$$=E(X^2)-2\mu^2+\mu^2$$
$$=E(X^2)-\mu^2$$

Gaussian Integral

The *Gaussian integral* is

$$\boxed{\int_{-\infty}^{\infty}e^{-x^2}dx=\sqrt{\pi}}\ .$$

Proof.

$$\left(\int_{-\infty}^{\infty}e^{-x^2}dx\right)^2=\left(\int_{-\infty}^{\infty}e^{-x^2}dx\right)\left(\int_{-\infty}^{\infty}e^{-y^2}dy\right)$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-(x^2+y^2)}dxdy$$
$$=\int_{0}^{2\pi}\int_{r=0}^{\infty}e^{-r^2}rdrd\theta \qquad (x=r\cos\theta,\ y=r\sin\theta)$$
$$=-\frac{1}{2}\int_{0}^{2\pi}\int_{r=0}^{\infty}e^{-r^2}d(-r^2)d\theta$$
$$=-\frac{1}{2}\left[e^{-r^2}\right]_{r=0}^{r=\infty}(2\pi-0)$$
$$=\pi$$
$$\Rightarrow\quad\int_{-\infty}^{\infty}e^{-x^2}dx=\sqrt{\pi}$$

Gaussian Distribution

Let $z=\dfrac{x-\mu}{\sigma}$ s.t. $\int_{-\infty}^{\infty} f(z)dx=1$ where $f(z)=ae^{-bz^2}$ is known as *Gaussian function*.

$$\Rightarrow \int_{-\infty}^{\infty} ae^{-b\left(\frac{x-\mu}{\sigma}\right)^2}dx=1 \Rightarrow a\int_{-\infty}^{\infty} e^{-\left(\sqrt{b}\frac{x-\mu}{\sigma}\right)^2} d\left(\sqrt{b}\frac{x-\mu}{\sigma}\right)=\frac{\sqrt{b}}{\sigma} \Rightarrow a=\frac{\sqrt{b}}{\sigma\sqrt{\pi}}$$

Take $b=\dfrac{1}{2}$, we have the probability density function, $f(z)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ as the *Gaussian distribution* (a.k.a. *normal distribution*.

Standard Normal Distribution

(1) Normal Distribution: $X \sim N(\mu, \sigma^2)$

A random variable X follows a normal distribution with mean μ and variance σ^2 is represented by $X \sim N(\mu, \sigma^2)$

(2) Standard Normal Distribution: $Z \sim N(0,1)$

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

$$E(Z)=E\left(\frac{X-\mu}{\sigma}\right)=\frac{1}{\sigma}E(X-\mu)=\frac{1}{\sigma}(\mu-\mu)=0$$

$$Var(Z)=Var\left(\frac{X-\mu}{\sigma}\right)=\frac{1}{\sigma^2}Var(X)=\frac{\sigma^2}{\sigma^2}=1$$

Sum / Difference of Normally Distributed Random Variables

If $X_1 \sim N(\mu_1, \sigma_1^{\,2})$ and $X_2 \sim N(\mu_2, \sigma_2^{\,2})$, then

- $(X_1 + X_2) \sim N(\mu_1 + \mu_2, \sigma_1^{\,2} + \sigma_2^{\,2})$
- $(X_1 - X_2) \sim N(\mu_1 - \mu_1, \sigma_1^{\,2} + \sigma_2^{\,2})$

Normal Distribution

Example.

(1) $X \sim N(10, 4)$. Find $P(X \geq 12)$. (Ans. 0.1587)

(2) $X \sim N(10, 4)$. Find $P(11 \geq X \geq 9.5)$. (Ans. 0.2902)

(3) $X \sim N(50, 100)$. If $P(X > a) = 0.04$, find the values of a. (Ans. $a = 67.5$)

(4) $X \sim N(50, 100)$. If $P(X < b) = 0.209$, find the values of b. (Ans. $b = 41.9$)