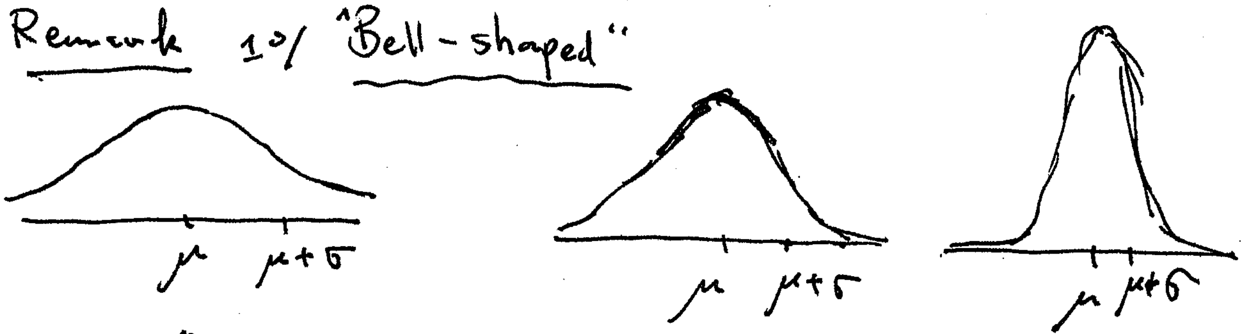
The Normal Distribution

The pdf of $X \sim N(\mu, \sigma^2)$:

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

Remark 1°/ "Bell-shaped"



↑ Normal density curves with parameters μ and σ^2.

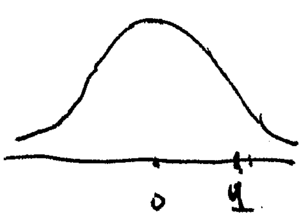
2°/

$$\int_{-\infty}^{\infty} f(x) = 1$$

Proof: $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ $\left(z = \frac{x-\mu}{\sigma} \right)$

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

$$= \int_{-\infty}^{\infty} \varphi(z)\,dz$$



standard normal curve

where $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ $\left(Z = \frac{X-\mu}{\sigma} \right)$

is the pdf of $Z \sim N(0,1)$.

The trick is to consider the square of the integral and convert it to a double integral:

p 2°/

$$\left(\int_{-\infty}^{\infty} \varphi(x)\,dx\right)\left(\int_{-\infty}^{\infty} \varphi(y)\,dy\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi(x)\varphi(y)\,dx\,dy$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy$$

(We can then use polar coordinates)
$\rho^2 = x^2+y^2$, $dx\,dy = \rho\,d\rho\,d\theta$

$$= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}\rho^2}\,\rho\,d\rho\,d\theta = \frac{1}{2\pi}\int_0^{2\pi}\left[-e^{-\frac{1}{2}\rho^2}\right]_0^{\infty} d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} 1\,d\theta = 1.$$

This establishes $\int_{-\infty}^{\infty} \varphi(x) = 1$ if taking the positive square root.

3°/ It is shown that for $X \sim N(\mu, \sigma^2)$

$$E(X) = \mu, \quad V(X) = \sigma^2$$

4°/ Let $X \sim N(\mu, \sigma^2)$. Then

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

$$Z = \frac{X-\mu}{\sigma} \qquad = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

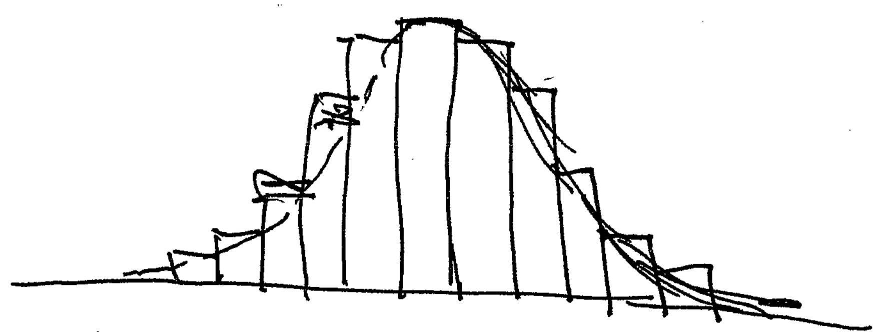
where $\Phi(z) = P(Z \le z)$, the shaded area, is tabulated in Appendix A.3

Approximating the binomial distribution

$$P(X \le x) = B(x; n, p) \simeq \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$$

provided that both $np \geq 10$ and $n(1-p) \geq 10$.

(de Moivre's result. it is a consequence of an important general result called the Central Limit theorem)



The (standard normal) mgf of $X \sim N(0, 1)$

$$M(t) = \int_{-\infty}^{\infty} e^{tx} \varphi(x)\,dx = \int_{-\infty}^{\infty} \exp\left(tx - \frac{x^2}{2}\right) dx$$

trick:

$$\left(tx - \frac{x^2}{2} = \frac{t^2}{2} - \frac{(x-t)^2}{2}\right)$$

$$= \int_{-\infty}^{\infty} \exp\left(\frac{t^2}{2}\right) \exp\left(-\frac{(x-t)^2}{2}\right) dx$$

$$= \exp\left(\frac{t^2}{2}\right) \int_{-\infty}^{\infty} \varphi(x-t)\,dx = \exp\left(\frac{t^2}{2}\right).$$

Expand the e^{tx} into its Taylor series in t and compare the result with Taylor series of $e^{\frac{t^2}{2}}$

$$M(t) = \int_{-\infty}^{\infty} \left\{ 1 + tx + \frac{(tx)^2}{2!} + \cdots + \frac{(tx)^n}{n!} + \cdots \right\} \varphi(x)\, dx$$

$$= 1 + \frac{t^2}{2} + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \cdots + \frac{1}{n!}\left(\frac{t^2}{2}\right)^n + \cdots$$

DeNote ~~that~~ the n^{th} moment of X by $m^{(n)}$

$$m^{(n)} = \frac{d^n}{dt^n}\left[M(t)\right] = \int_{-\infty}^{\infty} x^n \varphi(x)\, dx.$$

The above expansion becomes.

$$\sum_{n=0}^{\infty} \frac{m^{(n)}}{n!} t^n = \sum_{n=0}^{\infty} \frac{1}{2^n n!} \theta^{2n}$$

Then, from the uniqueness of power series expansion that the corresponding coefficients on both sides must be equal: thus for $n \geq 1$,

$$m^{(2n-1)} = 0 \qquad (\text{by symmetry too})$$

$$m^{(2n)} = \frac{(2n)!}{2^n n!}$$

Proposition: Let $\underline{X} \sim N(\mu, \sigma)$. Then

$$M_{\underline{X}}(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Proof.

trick:

$$M_{\underline{X}}(t) = E\left(e^{t\underline{X}}\right) = E\left(e^{t(\mu + \sigma Z)}\right)$$

$\left(\underline{X} = \mu + \sigma Z \text{ with } Z \sim N(0,1)\right)$

$$= e^{t\mu} E\left(e^{(t\sigma)Z}\right)$$

$$= e^{t\mu} M_Z(t\sigma) = e^{t\mu} e^{\frac{t^2\sigma^2}{2}}$$

$$= e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}$$

$$\left(\begin{aligned} R_X(t) &= \ln\left[M_{\underline{X}}(t)\right] \\ &= \mu t + \frac{\sigma^2 t^2}{2} \\ \mu &= R_X^{(1)}(0) = E(\underline{X}) \\ \sigma^2 &= R_X^{(2)}(0) = V(X) \end{aligned}\right)$$

Basic property of the normal family

Theorem: Let $\underline{X}_j \sim N(\mu_j, \sigma_j^2)$ be independent, $1 \le j \le n$. Then

$$\underline{X}_1 + \underline{X}_2 + \cdots + \underline{X}_n \sim N\left(\sum_{j=1}^n \mu_j, \sum_{j=1}^n \sigma_j^2\right).$$

Proof: it is sufficient to prove it for $n=2$, then by induction.

$$M_{\underline{X}_1+\underline{X}_2}(t) \overset{\text{product rule}}{=} M_{\underline{X}_1}(t) M_{\underline{X}_2}(t) = e^{\left(\mu_1 t + \frac{\sigma_1^2 t^2}{2}\right)} e^{\left(\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right)}$$

$$= \exp\left[(\mu_1 + \mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}\right]$$

which is the mgf of $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Hence, $\underline{X}_1 + \underline{X}_2$ has this normal distribution.