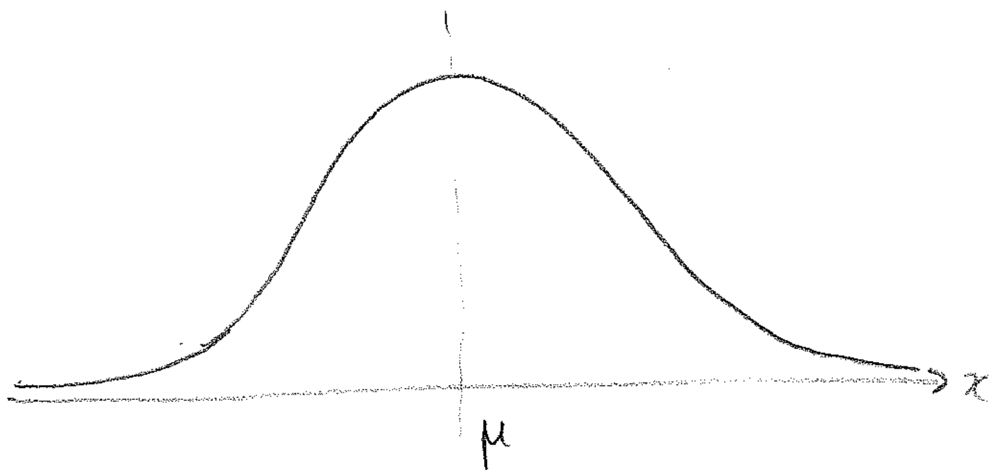
§ 6.5 The Normal Distribution

The most famous of them all!

Defn. A r.v. X has a normal (μ, σ) distribution iff its pdf can be written

$$n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty.$$

$\sigma > 0$.



'Bell Curve'.

An important special case is the normal (0,1) distribution, also called the <u>standard normal</u>

$$n(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}.$$

<u>Fact</u> $\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$

Using this it follows that the normal (0,1) distr has mass 1 and one can also see this for the general normal (μ, σ) since

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$z = \frac{x-\mu}{\sigma}$$

$$dz = \frac{dx}{\sigma}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \sigma \, e^{-z^2/2} dz$$

$$\sigma \, dz = dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi}$$

$$= 1.$$

Of course, we expect the normal (μ, σ) distr should have mean μ & var σ^2.

We show this via the MGF.

<u>Thm 6.6</u> The MGF of the normal (μ, σ) distr is given by

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

Pf By defn

$$M_X(t) = E(e^{Xt})$$

$$= \int_{-\infty}^{\infty} e^{xt} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left(-2\sigma^2 x t + (x-\mu)^2\right)} dx$$

We complete the square in the exponent. ie.

$$-2xt\sigma^2 + (x-\mu)^2 = -2xt\sigma^2 + x^2 - 2\mu x + \mu^2$$

$$= x^2 - 2(\mu + t\sigma^2)x + \mu^2$$

$$= x^2 - 2(\mu + t\sigma^2)x + (\mu + t\sigma^2)$$

$$- (\mu + t\sigma^2) + \mu^2$$

$$= (x - (\mu t + \sigma^2))^2$$

$$- (\cancel{\mu^2} + 2\mu t\sigma^2 + t^2\sigma^4) + \cancel{\mu^2}$$

$$= (x - (\mu t + \sigma^2))^2$$
$$- 2\mu t \sigma^2 - t^2\sigma^4$$

Then

$$M_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left[(x-(\mu + t\sigma^2))^2 - 2\mu t\sigma^2 - t^2\sigma^4\right]} dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left[x-(\mu + t\sigma^2)\right]^2 + \mu t + \frac{1}{2}\sigma^2 t^2} dx$$

$$= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{x-(\mu + t\sigma^2)}{\sigma}\right]^2} dx \right\}$$

The quantity inside the braces is the total mass of a normal with parameters $\mu + t\sigma^2$ and σ and so is just 1.

Thus

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \quad \text{as required.}$$

If we diff, we get

$$M_X'(t) = (\mu + \sigma^2 t)\, e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

$$M_X''(t) = \sigma^2 e^{\mu t + \frac{1}{2}\sigma^2 t^2} + (\mu + \sigma^2 t)^2 e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

$$= \left[(\mu + \sigma^2 t)^2 + \sigma^2\right] e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

If we then set $t = 0$ and apply Thm 4.9, we get

$$\mu_1' = M_X'(0) = \mu$$

$$\mu_2' \; M_X''(0) = \mu^2 + \sigma^2$$

Thus the normal (μ, σ) does indeed have mean μ and the variance is

$$\mu_2' - (\mu_1')^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

Thus. The mean and variance of the normal (μ, σ) distr are μ and σ^2 resp.

We can also go the other way and relate the normal (μ, σ) to the standard normal $(0,1)$.

Recall that if X is a r.v. with mean ν & var τ^2, then the r.v.

$$Y = a(X + b)$$

has mean $a(\nu + b)$

and variance $a^2 \tau^2$.

Thus is X has a normal (μ, σ) distr., then

$$Y = \frac{1}{\sigma}(X - \mu)$$

has mean $\frac{1}{\sigma}(\mu - \mu) = 0$

and variance $\frac{1}{\sigma^2} \cdot \sigma^2 = 1$.

We have proved

<u>Thm 6.7</u> If X is normal(μ, σ), then

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

is a standard normal r.v.

This result can be used to calculate probs. associated with normal r.v.'s.

The main drawback is that there is no nice formula for the CDF

$$F(x) = P(X \leq x)$$

of a normal, and we have to resort to tables. Let us denote by

$$\Phi(x) = P(Z \leq x),$$

the CDF of a std. normal $(0,1)$.

Ex. Let X be a normal (μ, σ) r.v.

Find the prob. the value of X is within 2 standard deviations of the mean.

Sol'n. We want

$$P(\mu - 2\sigma \le X \le \mu + 2\sigma)$$

$$= P(-2\sigma \le X - \mu \le 2\sigma)$$

$$= P\left(-2 \le \frac{X-\mu}{\sigma} \le 2\right)$$

$$= P(-2 \le Z \le 2)$$ where Z is a std. normal by thm 6.7.

$$= \Phi(2) - \Phi(-2)$$

$$\approx .95.$$