Note 14.

Section 3.3. Normal Distribution.

DEF X is said to have a normal distribution if its PDF is given by

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

In this case, we say that X is $N(\mu, \sigma^2)$.

PROP If X is $N(\mu, \sigma^2)$, then

$$E(X) = \mu, \quad Var(X) = \sigma^2, \quad M_X(t) = e^{\mu t + \frac{\sigma^2}{2}t^2}.$$

Pf) $M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$

Completing the square,

$$tx - \frac{(x-\mu)^2}{2\sigma^2} = \mu t + \frac{\sigma^2}{2}t^2 - \frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2},$$

and so,

$$M_X(t) = e^{\mu t + \frac{\sigma^2}{2}t^2} \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}}}_{\text{PDF of } N(\mu+\sigma^2 t,\, \sigma^2),\ \Rightarrow \text{ integrates } 1.} dx$$

$$= e^{\mu t + \frac{\sigma^2}{2}t^2}.$$

Then

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

and this gives

$$E(X) = M'(0) = \mu, \quad Var(X) = M''(0) - M'(0)^2 = \sigma^2. \quad \square$$

Ex If the PDF of X is

$$f_X(x) = \frac{1}{\sqrt{50\pi}} e^{-\frac{(x+3)^2}{50}},$$

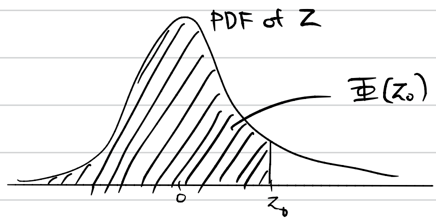
then X is $\mathcal{N}(-3, 5^2)$, and

$$M_X(t) = e^{-3t + \frac{25}{2}t^2},$$

and the converse is also true.

DEF
- If Z is $\mathcal{N}(0,1)$, then Z has a standard normal distribution.
- The CDF of a standard normal variable Z is denoted by

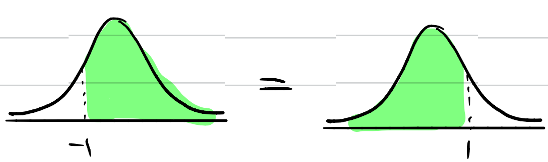
$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw.$$



Ex If Z is $\mathcal{N}(0,1)$, then
- $P(Z < 3) = P(Z \le 3) = \Phi(3)$,
- $P(Z > -1) = 1 - P(Z \le -1) = 1 - \Phi(-1)$.

  But since $\mathcal{N}(0,1)$ is symmetric about 0,

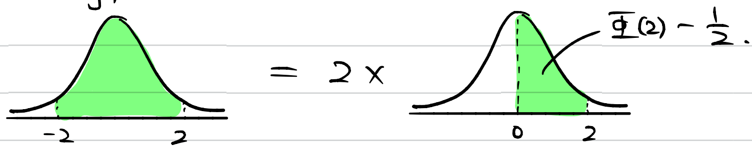
  $P(Z > -1) = P(-Z > -1) = P(Z < 1) = \Phi(1)$.



- $P(|Z|<2) = P(-2<Z<2)$
  $= P(Z<2) - P(Z\le -2)$
  $= \Phi(2) - \Phi(-2)$,

  but also,
  $= \Phi(2) - (1-\Phi(2))$
  $= 2\Phi(2) - 1.$

  Geometrically,



- $P(|Z|>3) = 2P(Z>3) = 2(1-\Phi(3)).$

<u>DEF</u> $z_\alpha := [100(1-\alpha)$ percentile$]$
$= [$ solution z of the equation $P(Z>z)=\alpha\ ].$

<u>THM</u> If X is $\mathcal{N}(\mu,\sigma^2)$, then $Z=\frac{X-\mu}{\sigma}$ is $\mathcal{N}(0,1)$.

<u>1st PF)</u> The CDF of Z is

$$F_Z(z) = P(Z\le z) = P\left(\frac{X-\mu}{\sigma}\le z\right)$$
$$= P(X\le \mu+\sigma z) = F_X(\mu+\sigma z).$$

$$\Rightarrow f_Z(z) = \frac{d}{dz}F_Z(z) = \frac{d}{dz}F_X(\mu+\sigma z)$$
$$= \sigma\cdot f_X(\mu+\sigma z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}.$$

This is the PDF of $\mathcal{N}(0,1)$. $\square$

2nd pf) The MGF of Z is

$$M_Z(t) = E[e^{tZ}] = E[e^{t\cdot\frac{X-\mu}{\sigma}}]$$
$$= e^{-\mu t/\sigma} E[e^{(t/\sigma)X}]$$
$$= e^{-\mu t/\sigma} M_X(t/\sigma)$$
$$= e^{-\mu t/\sigma} e^{\mu(t/\sigma)+\frac{\sigma^2}{2}(t/\sigma)^2}$$
$$= e^{\frac{t^2}{2}}.$$

This is the MGF of $N(0,1)$. $\square$

Rmk) In this context,

- $Z = (X-\mu)/\sigma$ is called the standard score,
- Transforming X to Z is called standardization.

Ex If X is $N(4, 5^2)$, then

$$P(3 < X < 6) = P\left(\frac{3-4}{5} < \frac{X-4}{5} < \frac{6-4}{5}\right)$$
$$= P\left(-\frac{1}{5} < Z < \frac{2}{5}\right)$$
$$= \Phi\left(\frac{2}{5}\right) - \Phi\left(-\frac{1}{5}\right). \quad \square$$

Ex If the volume X of a certain type of canned beer is distributed according to a normal dist. with mean 12.1 oz, then what is the range of σ s.t.

$$P(X > 12) \geq 0.99?$$

<u>Sol)</u>

$$P(X > 12) = P\left(\frac{X-12.1}{\sigma} > \frac{12-12.1}{\sigma}\right)$$

$$= P\left(Z > -\frac{1}{10\sigma}\right)$$

$$= 1 - P\left(Z > \frac{1}{10\sigma}\right).$$

So,

$$0.99 \leq P(X>12) \iff P\left(Z > \frac{1}{10\sigma}\right) \leq 0.01$$

and this implies

$$\frac{1}{10\sigma} \geq z_{0.01} = 2.326$$

$$\iff \sigma \leq \frac{1}{23.26} = 0.043.$$

$\square$