The Normal Distribution

One of the most important distributions in statistical applications. For now we just introduce it to gain familiarity.

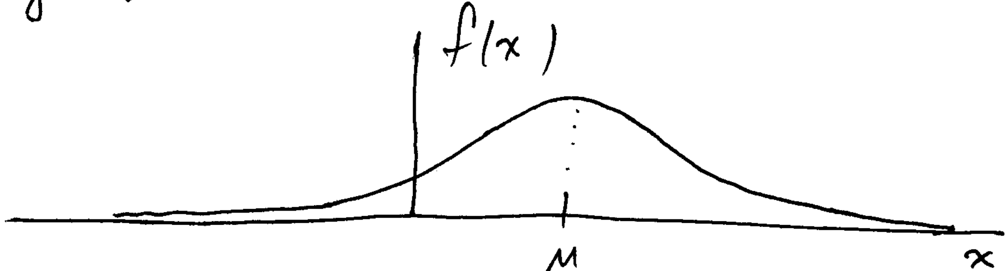
<u>Def</u> Let $\mu \in \mathbb{R}$, $\sigma^2 > 0$ be parameters. X has a <u>normal distribution</u> if

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

<u>Notation</u>: $X \sim N(\mu, \sigma^2)$

<u>Key fact</u>:
$$E[X] = \mu$$
$$E[(X-\mu)^2] = \sigma^2$$

which justifies the use of these parameter names.



<u>Normalization</u>

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

Proof: see book.

$\boxed{MGF}$ $$M(t) = E\left[e^{Xt}\right] = \int_{-\infty}^{\infty} e^{xt} f(x)\, dx$$

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

<u>Side calculation</u>: $e^{xt} e^{-(x-\mu)^2/2\sigma^2}$

$$= \exp\left[xt - \frac{(x^2 - 2x\mu + \mu^2)}{2\sigma^2} \right]$$

$$= \exp\left[\frac{2xt\sigma^2 - x^2 + 2x\mu - \mu^2}{2\sigma^2} \right]$$

$$= \exp\left[-\frac{(x^2 - 2(\mu + \sigma^2 t)x + \mu^2)}{2\sigma^2} \right]$$

Complete the square :

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

$$\underbrace{x^2 - 2(\mu+\sigma^2 t)x + (\mu+\sigma^2 t)^2}_{(x - (\mu+\sigma^2 t))^2} \underbrace{- (\mu+\sigma^2 t)^2 + \mu^2}_{-2\mu\sigma^2 t - \sigma^4 t^2}$$

Thus

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

$$= \exp\left(\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2} \right) \cdot \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[- \frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2} \right]}_{=1}$$

(pdf of $N(\mu+\sigma^2 t, \sigma^2)$).

$$= \exp\left(\mu t + \frac{\sigma^2 t^2}{2} \right). \qquad \text{for all } t \in \mathbb{R}$$

$$M'(t) = (\mu + \sigma^2 t)\exp\left(\mu t + \frac{\sigma^2 t}{2}\right)$$

$$M'(0) = \mu \quad \checkmark$$

$$M''(t) = \sigma^2 \exp\left(\mu t + \frac{\sigma^2 t}{2}\right) + (\mu + \sigma^2 t)^2 \exp\left(\mu t + \frac{\sigma^2 t}{2}\right)$$

$$M''(0) = \sigma^2 + \mu^2$$

$$\Rightarrow \mathrm{Var}(X) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \quad \checkmark$$

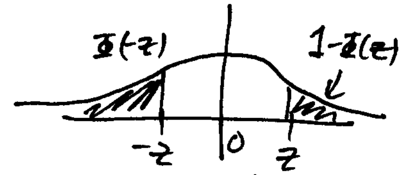
<u>Standard Normal</u>

$$Z \sim N(0,1)$$

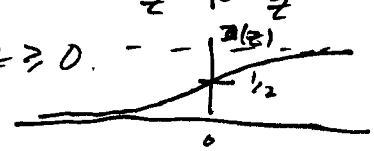
CDF:

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}\, dy$$

No closed form expression.



<u>Question</u> What if we only know $\Phi(z)$ for $z \geq 0$.



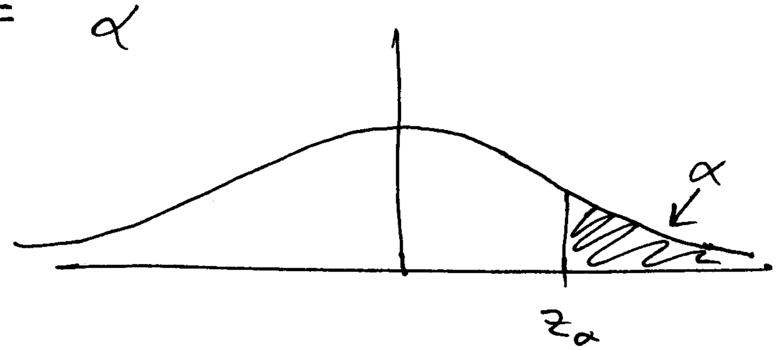
What is $\Phi(-z)$? <u>Ans</u> $\Phi(-z) = 1 - \Phi(z)$

Upper percentiles

Let $\alpha \in (0, \frac{1}{2})$

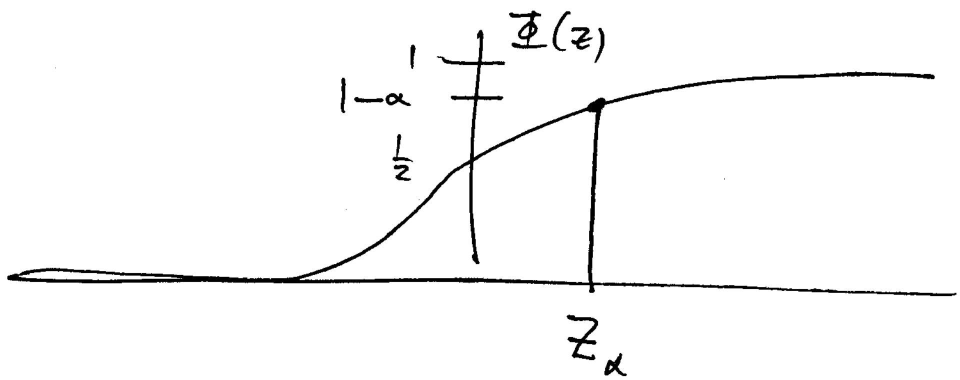
Define $z_\alpha = \pi_{1-\alpha}$ to be the points such that

$$P(Z \geq z_\alpha) = \alpha$$



In terms of the CDF,

$$1 - \Phi(z_\alpha) = \alpha \implies z_\alpha = \Phi^{-1}(1-\alpha)$$



<u>Exercise</u>: Express $z_{1-\alpha}$ in terms of z_α.

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

<u>Proof</u>: We will show $P(Z \le z) = \Phi(z)$.

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

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

Let $y = \frac{x-\mu}{\sigma}$. Then $dx = \sigma\, dy$

Lower limit : $y = \frac{-\infty - \mu}{\sigma} = -\infty$

Upper " : $y = \frac{(\sigma z + \mu) - \mu}{\sigma} = z$

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

Since the CDF uniquely determines the distribution, we conclude $Z \sim N(0,1)$.

Example Suppose $X \sim N(3, 5)$.

What is $P(X \le 4)$?

$$
\begin{aligned}
P(X \le 4) &= P(X - 3 \le 4 - 3) \\
&= P(X - 3 \le 1) \\
&= P\left(\frac{X-3}{\sqrt{5}} \le \frac{1}{\sqrt{5}}\right) \\
&= \Phi\left(\frac{1}{\sqrt{5}}\right) \\
&= \Phi(.4472) \\
&= .6726
\end{aligned}
$$

<u>Theorem</u> If $X \sim N(\mu, \sigma^2)$ and $V = \left(\frac{X-\mu}{\sigma}\right)^2$,

then $V \sim \chi^2(1)$.

<u>Proof</u>:

Recall the pdf of a $\chi^2(r)$ random variable is

$$f(x) = \frac{1}{\Gamma(\frac{r}{2})\, 2^{r/2}}\, x^{\frac{r}{2}-1} e^{-x/2} \quad , x \geq 0$$

If $r=1$,

$$f(x) = \frac{1}{\Gamma(\frac{1}{2})\sqrt{2}\,\sqrt{x}}\, e^{-x/2}.$$

<u>Fact</u>: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$\Longrightarrow f(x) = \frac{1}{\sqrt{2\pi x}}\, e^{-x/2} \quad , x \geq 0.$$

Let $F(v)$ denote the CDF of V. It suffices to show

$$F(v) = \int_0^v \frac{1}{\sqrt{2\pi y}} e^{-y/2}\, dy$$

Let $Z = \frac{X-\mu}{\sigma^2}$, so that $V = Z^2$.

Then

$$\begin{aligned}
F(v) &= P(V \le v) \\
&= P(Z^2 \le v) \\
&= P(-\sqrt{v} \le Z \le \sqrt{v}) \\
&= \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz
\end{aligned}$$

$$\left[\text{Let } y = z^2 \Rightarrow dy = 2z\, dz \Rightarrow dz = \frac{dy}{2\sqrt{y}}\right]$$

$$= \int_0^y \frac{1}{\sqrt{2\pi y}} e^{-y/2}\, dy$$

Remark: Since above must integrate to 1, can use to prove $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.