CONTINUOUS RANDOM VARIABLES

The normal distribution

The normal distribution

- The normal (Gaussian) distribution is certainly the most commonly used continuous distribution as a model in statistics.
- Many real-world data sets seem to follow the normal ("bell shaped") distribution.
- The normal distribution is defined over the entire real line.

The normal distribution

- Definition 5: A random variable Y has a NORMAL DISTRIBUTION with mean μ and variance σ^2 if its pdf is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \qquad -\infty < y < \infty$$

- Shorthand notation: $Y \sim N(\mu, \sigma^2)$. (Note: must have $\sigma > 0$.)

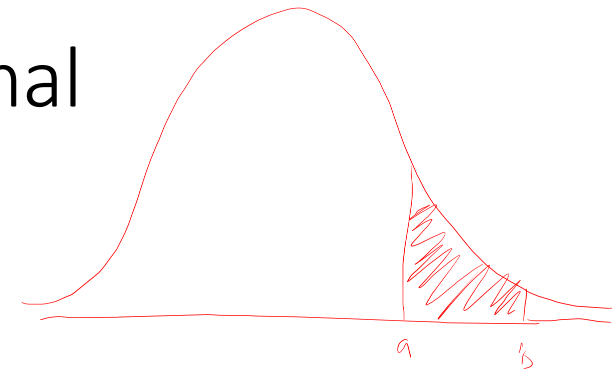
The normal distribution

- The normal distribution is symmetric about μ.
- It has inflection points $\mu - \sigma$ and $\mu + \sigma$.
- The normal pdf is positive everywhere, but
$$\lim_{y \to -\infty} f(y) = \lim_{y \to \infty} f(y) = 0$$
- Because $f(y)$ is a pdf,

$$\int_{-\infty}^{\infty} f(y) dy = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}\, dy = 1$$

(This is *really hard* to prove!)

Calculating probabilities with the normal distribution



- If $Y \sim N(\mu, \sigma^2)$ then

$$P(a < Y < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}\, dy$$

- Problem: Nobody knows how to solve this integral!
- Solution: Numerical approximation!
  - Table 4, Appendix 3
  - `pnorm()` function in `R`
  - Online calculators
- Assignment: figure out how to use (at least) one of these.
- Useful fact: If $Y \sim N(\mu, \sigma^2)$ then $\frac{Y-\mu}{\sigma} \sim N(0,1)$

The standard normal distribution

- The *parameters* of the $N(\mu, \sigma^2)$ distribution are μ and σ^2.
- If a normal random variable Z has $\mu = 0$ and $\sigma^2 = 1$ (i.e., $Z \sim N(0,1)$) then we say that Z has the standard normal distribution.

Moment generating function of the standard normal distribution

- Theorem 4: If Z~$N(0,1)$ then Z has mgf

$$m(t) = E[e^{tZ}] = \boxed{e^{\frac{1}{2}t^2}}$$

$$m(t) = E[e^{tz}] = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$e^{tz} e^{-\frac{1}{2}z^2} = e^{-\frac{1}{2}z^2 + zt - \frac{1}{2}t^2 + \frac{1}{2}t^2} = e^{-\frac{1}{2}(z-t)^2} e^{\frac{1}{2}t^2}$$

$$= \int_{-\infty}^{\infty} e^{\frac{1}{2}t^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz = e^{\frac{1}{2}t^2} \underbrace{\int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2}}_{\text{pdf of } N(t,1)\,!!} dz}_{=1}$$

$$= e^{\frac{1}{2}t^2}$$

Moment generating function of the normal distribution

- Theorem 5: If $Y \sim N(\mu, \sigma^2)$ then Y has mgf

$$m(t) = E[e^{tY}] = e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$$

$$m(t) = E[e^{tY}] = \int_{-\infty}^{\infty} e^{ty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}(y-\mu)^2/\sigma^2}\, dy$$

substitute $z = \frac{y-\mu}{\sigma} \Leftrightarrow y = z\sigma + \mu$ $\quad dy = \sigma\, dz$

$$= \int_{-\infty}^{\infty} e^{\underbrace{t\sigma z + t\mu}_{\text{no } z}} \frac{1}{\sqrt{2\pi}\,\cancel{\sigma}} e^{-\frac{1}{2}z^2} \cancel{\sigma}\, dz$$

$$= e^{t\mu} \underbrace{\int_{-\infty}^{\infty} e^{tz\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz}$$

mgf of $N(0,1)$ evaluated at $t\sigma$

$$= e^{t\mu} \cdot e^{\frac{1}{2}(t\sigma)^2} = \boxed{e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}}$$

Mean and variance of the normal distribution

- Theorem 6: If $Y \sim N(\mu, \sigma^2)$ then

$$E[Y] = \mu \text{ and } Var(Y) = \sigma^2.$$

Take $Z \sim N(0,1)$ want to show $E[Z] = 0$

$$E[Z] = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} z e^{-\frac{z^2}{2}} dz + \frac{1}{2\pi} \int_{0}^{\infty} z e^{-\frac{z^2}{2}} dz$$

substitute $u = -\frac{z^2}{2}$

$du = -z\,dz$

$z = 0 \Leftrightarrow u = 0$

"$z = \infty$" $\Leftrightarrow$ "$u = -\infty$"

"$z = -\infty$" $\Leftrightarrow$ "$u = -\infty$"

$$= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} \cancel{e^{u} du} - \int_{-\infty}^{0} \cancel{e^{u} du} \right) = 0$$

If $Z \sim N(0,1)$ $Y = \sigma Z + \mu \sim N(\mu, \sigma^2)$

$$E(Y) = E(\sigma Z + \mu) = \sigma \underbrace{E[Z]}_{0} + \mu = \mu$$