Normal – Normal

Given : $x_i \mid \mu \sim N(\mu, \sigma^2)$ i.i.d

$\mu \sim N(\beta, \tau^2)$ (prior)

Then : $\mu \mid x_1, \dots, x_n = \mu \mid \bar{x} \sim N\left(\dfrac{\beta \sigma^2/n + \bar{x}\tau^2}{\sigma^2/n + \tau^2}, \dfrac{\tau^2 \sigma^2/n}{\sigma^2/n + \tau^2} \right)$

where $\bar{x} = \sum_{i=1}^{n} x_i / n$ (average)

posterior for μ

Conjugate prior

Application to a classical problem

mean of two groups

Given: $x_i \mid \mu_x \sim N(\mu_x, \sigma_x^2)$ n i.i.d samples

$y_i \mid \mu_y \sim N(\mu_y, \sigma_y^2)$ m i.i.d samples

Decide if $\mu_x = \mu_y$. Classical approach: Construct $z = \dfrac{\bar{x} - \bar{y}}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}}$

What do we know?

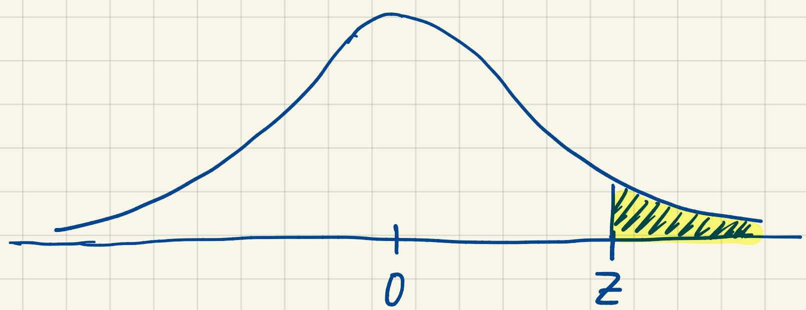
$\bar{x} \sim N(\mu_x, \sigma_x^2/n)$ $\bar{y} \sim N(\mu_y, \sigma_y^2/m)$ $\bar{x} - \bar{y} \sim N\left(\mu_x - \mu_y, \dfrac{\sigma_x^2}{n} + \dfrac{\sigma_y^2}{m}\right)$

[linear combination]

$$z = \frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}} \sim N(0,1)$$

If $\mu_x = \mu_y \Rightarrow \mu_x - \mu_y = 0$, and

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}} \sim N(0,1)$$



Find $\Phi(-|z|)$

Called P-value

If P-value is small, reject that $\mu_x = \mu_y$

(because z is not likely to show up either too small or too big)

Note: P-value is not $P(\mu_x = \mu_y)$

Bayesian approach

$$\mu_x \mid \bar{x} \sim N(m_1, V_1^2)$$

$$\mu_y \mid \bar{y} \sim N(m_2, V_2^2)$$

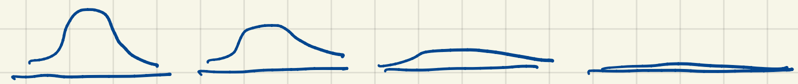
$$\mu_x - \mu_y \mid \bar{x}, \bar{y} \sim N(\underbrace{m_1 - m_2}_{m}, \underbrace{V_1^2 + V_2^2}_{V^2})$$

We have the density of $\mu_x - \mu_y$ conditioned on $\bar{x}$ and $\bar{y}$.

For instance, we can find $P(\mu_x - \mu_y > b \mid \bar{x}, \bar{y}) = \Phi\left(\frac{m-b}{V}\right)$

This is now a little more informative since we can choose b as the threshold for what it means $\mu_x \neq \mu_y$.

Another way to look at this is to make $\tau^2 \to \infty$

in $\mu \sim N(\beta, \tau^2)$. Then 

$$\mu_x \mid \bar{x} \sim N\left(\bar{x}, \sigma_x^2/n\right) \qquad \text{[same result as above]}$$

$$\mu_y \mid \bar{y} \sim N\left(\bar{y}, \sigma_y^2/m\right)$$

$$\mu_x - \mu_y \mid \bar{x}, \bar{y} \sim N\Big(\underbrace{\bar{x} - \bar{y}}_{m}, \underbrace{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}_{v^2}\Big)$$

$$P(\mu_x - \mu_y > b \mid \bar{x}, \bar{y}) = \Phi\left(\frac{m-b}{v}\right) = \Phi\left(\frac{\bar{x} - \bar{y} - b}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}\right)$$

When $b=0$, $P(\mu_x > \mu_y \mid \bar{x}, \bar{y}) = \Phi\left(\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}\right) = \Phi(z)$

Recall, I defined P-value as $\Phi(-|z|)$.

$$\underset{\Phi(z)}{P(\mu_x > \mu_y \mid \bar{x}, \bar{y})} = \begin{cases} 1 - \text{P-value} & z \geq 0 \\ \text{P-value} & z \leq 0 \end{cases}$$

Interpretation: P-value small and $z \geq 0 \Rightarrow P(\mu_x > \mu_y \mid \bar{x}, \bar{y})$ high

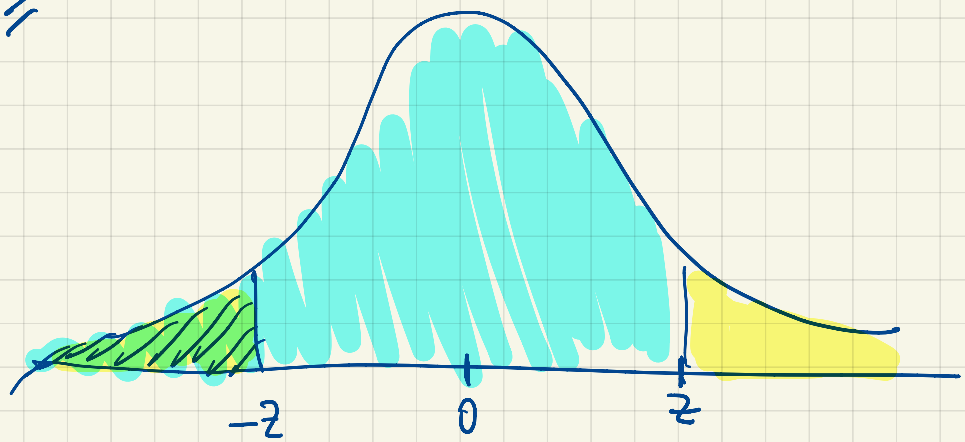
P-value small and $z \leq 0 \Rightarrow P(\mu_y > \mu_x \mid \bar{x}, \bar{y})$ high

P-value moderate $\Rightarrow$ $P(\mu_x > \mu_y \mid \bar{x}, \bar{y})$, $P(\mu_y > \mu_x \mid \bar{x}, \bar{y})$ not high

Better interpretation of what P-value is.

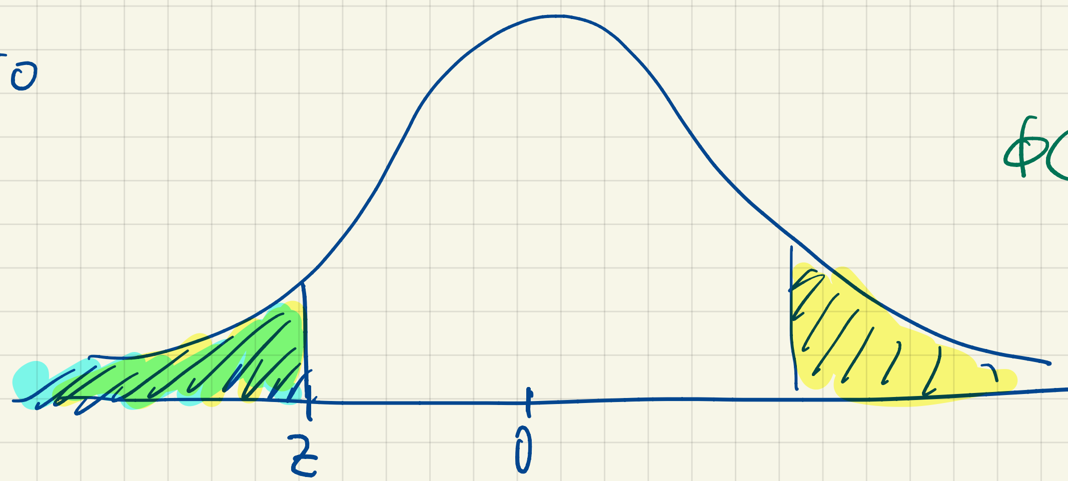
$P_{value} = \Phi(-|z|)$

$z \geqslant 0$



$\phi(z) = 1 - P_{value}$

$z \leqslant 0$



$\phi(z) = P\text{-value}$

What does this all mean?

- Bayesian approach offers a more general framework and with a special choice for the prior, it boils down to the classical approach.

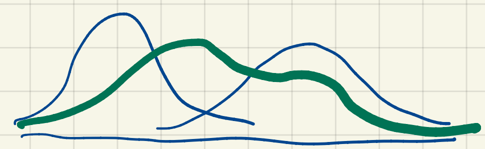
- This choice of prior is usually <u>improper</u>. We call it <u>reference</u> prior, because it brings us back to the non-Bayesian setting.

reference prior was uniform from $-\infty$ to $+\infty$, which meant I know nothing

- Another advantage of the Bayesian approach is that we can "mix" priors. ("not sure")

Mixed Prior "Mixed feelings"

Assume prior is $f(\theta) = \alpha \underbrace{g(\theta)}_{\text{prior 1}} + \beta \underbrace{h(\theta)}_{\text{prior 2}}$

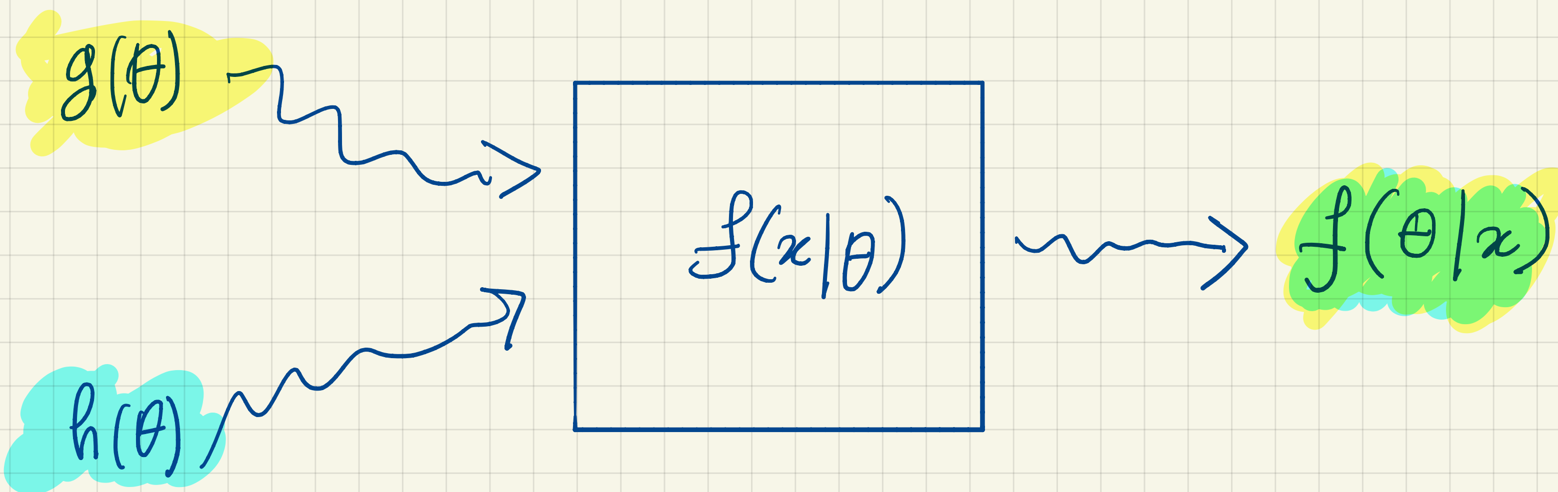


where $\alpha + \beta = 1$.

Observe:

$$\int f(\theta)\, d\theta = \int [\alpha g(\theta) + \beta h(\theta)]\, d\theta$$

$$= \alpha \int g(\theta)\, d\theta + \beta \int h(\theta)\, d\theta = \alpha + \beta = 1.$$



$$f(\theta|x) = \frac{f(x|\theta)\left[\alpha g(\theta) + \beta h(\theta)\right]}{f(x)}$$

$$= \alpha \frac{f(x|\theta)\, g(\theta)}{f(x)} + \beta \frac{f(x|\theta)\, h(\theta)}{f(x)}$$

Now $f(x|\theta)\, g(\theta) = g(\theta|x)\, g(x) = g(\theta|x) \int f(x|\theta) g(\theta)\, d\theta$

Similarly $f(x|\theta)\, h(\theta) = h(\theta|x) \int f(x|\theta)\, h(\theta)\, d\theta$

$$f(\theta|x) = \frac{\alpha \int f(x|\theta) g(\theta) d\theta}{f(x)}\, g(\theta|x) + \frac{\beta \int f(x|\theta) h(\theta) d\theta}{f(x)}\, h(\theta|x)$$

$$= \alpha(x)\, g(\theta|x) + \beta(x)\, h(\theta|x)$$

Where $\alpha(x) + \beta(x) = 1$ and $\dfrac{\alpha(x)}{\beta(x)} = \dfrac{\alpha}{\beta} \dfrac{\int f(x|\theta) g(\theta) d\theta}{\int f(x|\theta)\, h(\theta)\, d\theta}$

Example: $x \mid \mu \sim N(\mu, 1)$

$$\mu \sim 0.5\, N(0,1) + 0.5\, N(0,2)$$

- if $\mu \sim N(0,1)$ then $\mu \mid x \sim N\left(\frac{x}{2}, \frac{1}{2}\right)$

- if $\mu \sim N(0,2)$ then $\mu \mid x \sim N\left(\frac{2x}{3}, \frac{2}{3}\right)$

Assume we are observing $x=0$, then

$\mu \mid x=0$ is a mixture of $N(0,\frac{1}{2})$ and $N(0,\frac{2}{3})$

What's the exact mixture?

$$\frac{\alpha(0)}{\beta(0)} = \frac{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2}}\, d\mu}{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2}} \cdot \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{\mu^2}{4}}\, d\mu} \qquad \alpha(0)+\beta(0)=1$$

$$\frac{\alpha(0)}{\beta(0)} = \frac{\sqrt{2}\int \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2(1/2)}} d\mu}{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2(2/3)}} d\mu}$$

and $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2a}} d\mu = \sqrt{a}$

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi a}} e^{-\frac{\mu^2}{2a}} d\mu = 1$$

$$= \frac{\sqrt{2}\sqrt{1/2}}{\sqrt{2/3}} = \sqrt{3/2}$$

$$\alpha(0) = \sqrt{\tfrac{3}{2}}\,\underbrace{\left(1-\alpha(0)\right)}_{\beta(0)} \Rightarrow \alpha(0) = \frac{\sqrt{3/2}}{1+\sqrt{3/2}}$$

$$\beta(0) = \frac{1}{1+\sqrt{3/2}}$$