Week16-notes

Monday, 3 May 2021 12:09

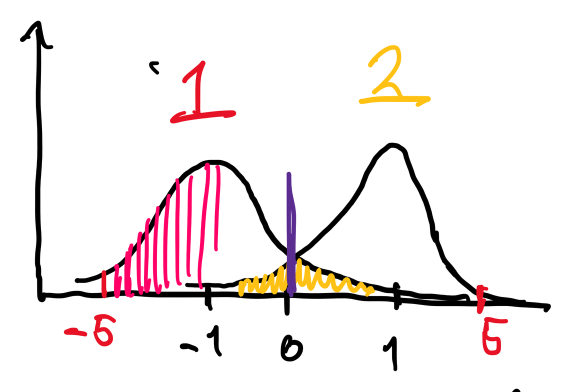
- Mandatory assignment 2 corrected.
- 2nd to last plenary session.
- Skip Extra 14, (done it before)
- Discuss Extra 13.
- Go through Extra 15
- Point out error in solution Exam 2019.
- End early today.

Extra 15

$Pr(Y=1) = Pr(Y=2) = 0.5$

$X \mid Y=k \sim N(\mu_k, 1)$

where $\mu_1 = -1$, $\mu_2 = 1$.



$X \mid Y$	$Y=1$	$Y=2$	$Y=3$
$X=1$	0.4	0.05	0.05
$X=2$	0.2	0.1	0.2

$\sum_X \sum_Y Pr(Y, X) = 1$

$Pr(X)$

$Pr(Y=1) = 0.6$

We want to comput.

$$Pr(Y=k \mid X) = \frac{Pr(X \mid Y=k)\, Pr(Y=k)}{\boxed{Pr(X)}} \quad (*)$$

Bayes rule

Need to find $Pr(X)$.

$Pr(Y, X)$

$$Pr(X) = \sum_{k=1}^{2} Pr(X \mid Y=k) \underbrace{Pr(Y=k)}_{0.5}$$

$$= \frac{1}{2} \sum_{k=1}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_k)^2}$$

$$= \frac{1}{2\sqrt{2\pi}} \left(e^{-\frac{1}{2}(x+1)^2} + e^{-\frac{1}{2}(x-1)^2} \right)$$

Insert all of this into $(*)$, we get

$$Pr(Y=k \mid X) = \frac{\cancel{\frac{1}{2}} \cdot \cancel{\frac{1}{\sqrt{2\pi}}}\, e^{-\frac{1}{2}(x-\mu_k)^2}}{\frac{1}{2\sqrt{2\pi}} \left(e^{-\frac{1}{2}(x+1)^2} + e^{-\frac{1}{2}(x-1)^2} \right)}$$

This means that

Bayes classifier: $\arg\max_k Pr(Y=k \mid X)$

$$= \arg\max_k \left\{ \frac{e^{-\frac{1}{2}(x-\mu_k)^2}}{e^{-\frac{1}{2}(x+1)^2} + e^{-\frac{1}{2}(x-1)^2}} \;\middle|\; X \right\}$$

Simplify it further.

Examine the decision boundary.

$Pr(Y=1 \mid X) > Pr(Y=2 \mid X)$.

$$e^{-\frac{1}{2}(x+1)^2} > e^{-\frac{1}{2}(x-1)^2}$$

$$-\frac{1}{2}(x+1)^2 > -\frac{1}{2}(x-1)^2$$

$$-(x+1)^2 > -(x-1)^2$$

$$-[x^2+2x+1] > -[x^2-2x+1]$$

$$\cancel{-x^2} - 2x \cancel{-1} > \cancel{-x^2} + 2x \cancel{-1}$$

$$0 > 4x$$

$$0 > x.$$

So, we have the following Bayes classifier

$$k^{Bayes} = \arg\min_k \{1 - Pr(Y=k \mid X)\}$$

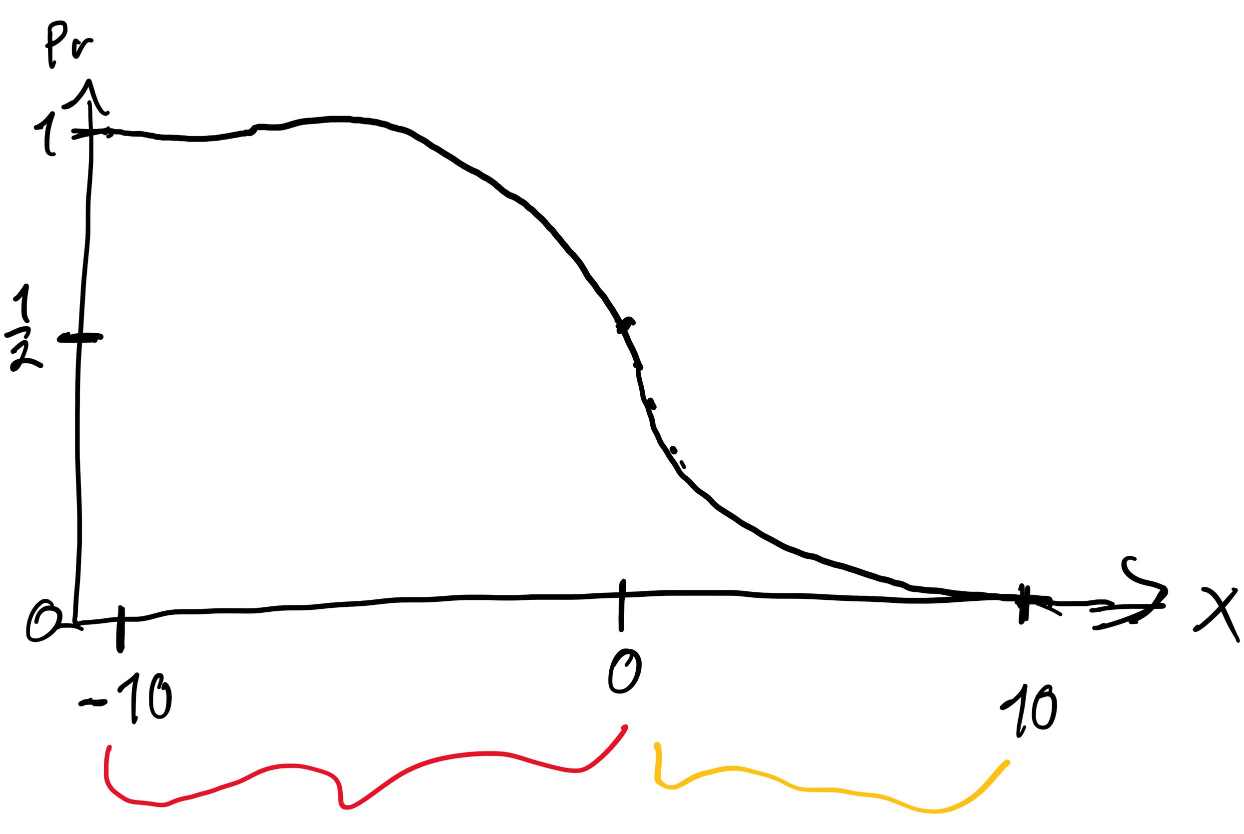
$$= \arg\max_k Pr(Y=k \mid X)$$

$$= \begin{cases} 1 & \text{if } x < 0 \\ 2 & \text{if } x \geq 0 \end{cases}$$

b)

Plot $Pr(Y=1 \mid X=x)$

$$= \frac{e^{-\frac{1}{2}(x+1)^2}}{e^{-\frac{1}{2}(x+1)^2} + e^{-\frac{1}{2}(x-1)^2}}$$



c)

Define, marginal dist.

$$f_X(x) = \sum_{k=1}^{2} Pr(X \mid Y=k)\, Pr(Y=k)$$

$$= \frac{1}{2\sqrt{2\pi}} \left(e^{-\frac{1}{2}(x+1)^2} + e^{-\frac{1}{2}(x-1)^2} \right)$$

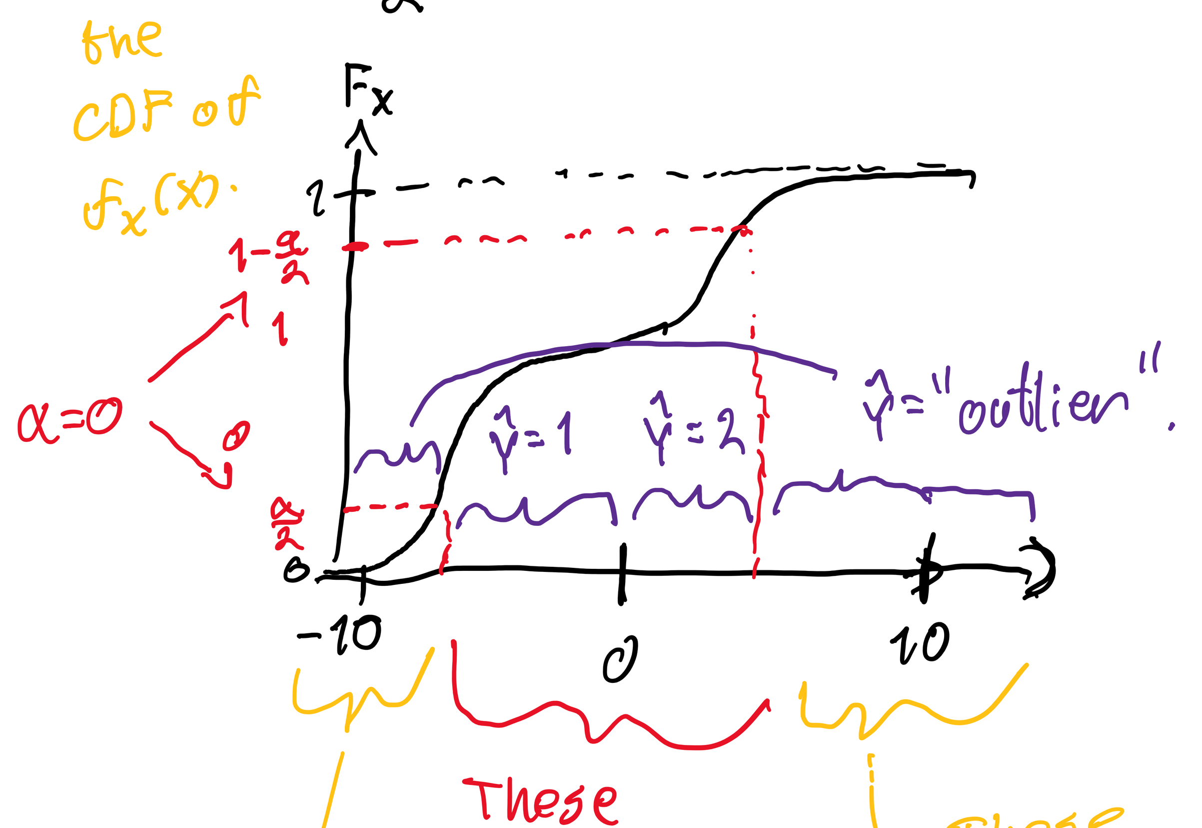
this is a mixture normal. $f_X(x) = \frac{1}{2}N(-1,1) + \frac{1}{2}N(1,1)$. which is a legit pdf.

Furthermore, let

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du$$

This is the CDF of $f_X(x)$.

$$= \frac{1}{2}\Phi(x+1) + \frac{1}{2}\Phi(x-1)$$



These values are plausible

These are considerd outliers,

We get the following Null hypothesis testing.

Reject H_0 if $F_X(x) < \frac{\alpha}{2}$ or if $F_X(x) > 1 - \frac{\alpha}{2}$.

d) and onwards

Look at code.