

## LECTURE 10: Conditioning on a random variable; Independence; Bayes' rule

- Conditioning  $X$  on  $Y$ 
  - Total probability theorem
  - Total expectation theorem
- Independence
  - independent normals
- A comprehensive example
- Four variants of the Bayes rule

## Conditional PDFs, given another r.v.

$$p_{X|Y}(x | y) = \mathbf{P}(X = x | Y = y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}, \quad \text{if } p_Y(y) > 0$$

$p_{X,Y}(x, y)$	$f_{X,Y}(x, y)$
$p_{X A}(x)$	$f_{X A}(x)$
$p_{X Y}(x   y)$	$f_{X Y}(x   y)$

**Definition:**  $f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$  if  $f_Y(y) > 0$

$$\mathbf{P}(x \leq X \leq x + \delta | A) \approx f_{X|A}(x) \cdot \delta, \quad \text{where } \mathbf{P}(A) > 0$$

$\underbrace{\hspace{1.5cm}}_{Y=\gamma} \downarrow Y \approx \gamma$

$$\mathbf{P}(x \leq X \leq x + \delta | y \leq Y \leq y + \epsilon) \approx \frac{f_{X,Y}(x, \gamma) \delta}{f_Y(\gamma)} = f_{X|Y}(x | \gamma) \delta$$

Definition:  $\mathbf{P}(X \in A | Y = \underset{\bullet}{y}) = \int_A f_{X|Y}(x | y) dx$

## Comments on conditional PDFs

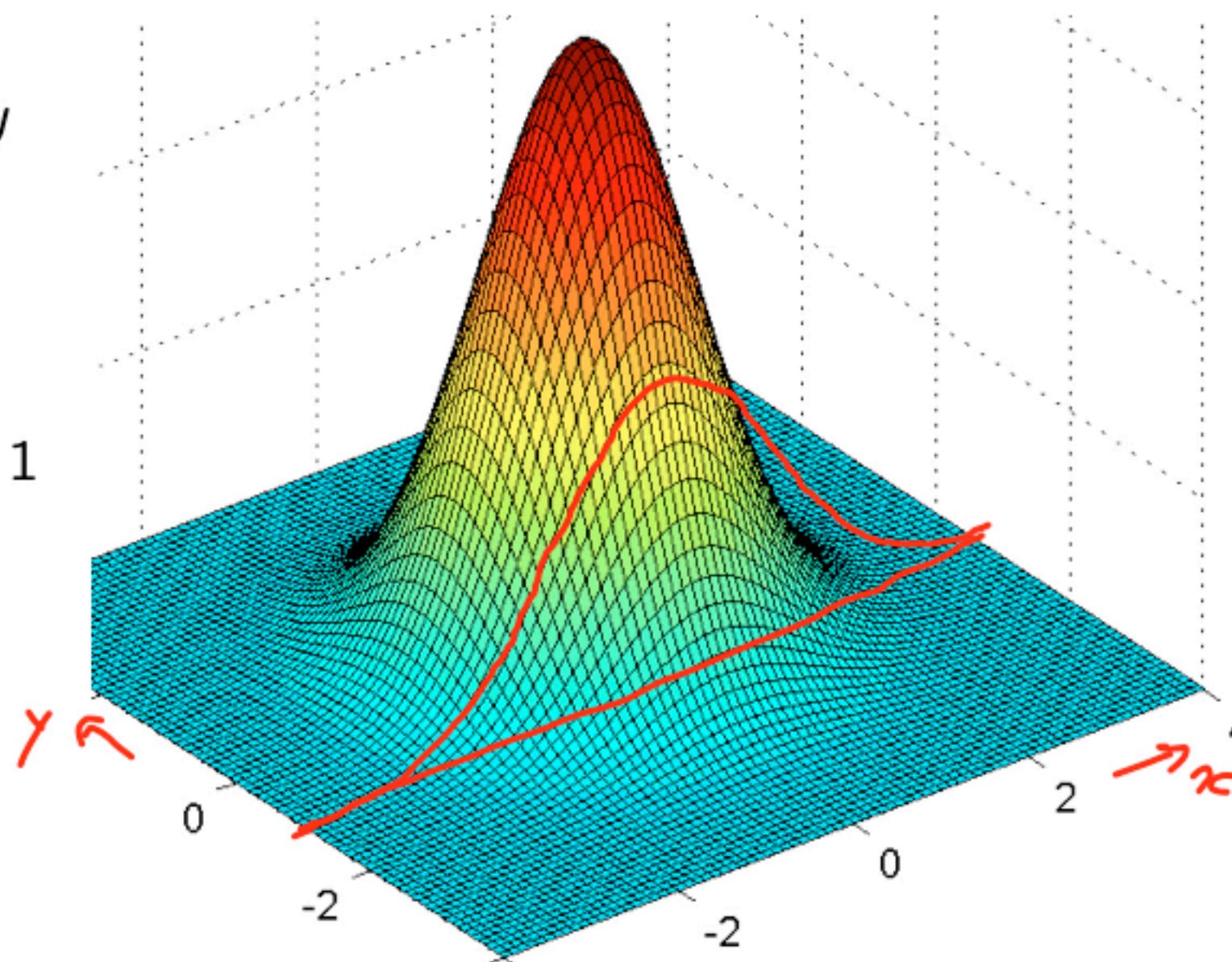
$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad \bullet \quad f_{X|Y}(x | y) \geq 0$$

- Think of value of  $Y$  as fixed at some  $y$   
shape of  $f_{X|Y}(\cdot | y)$ : slice of the joint

$$\bullet \quad \int_{-\infty}^{\infty} f_{X|Y}(x | y) dx = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x, y) dx}{f_Y(y)} = 1$$

- Multiplication rule:

$$\begin{aligned} f_{X,Y}(x, y) &= f_Y(y) \cdot f_{X|Y}(x | y) \\ &= f_X(x) \cdot f_{Y|X}(y | x) \end{aligned}$$





## Total probability and expectation theorems

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x|y)$$

$$f_X(x) = \int_{-\infty}^{\infty} \overbrace{f_Y(y) f_{X|Y}(x|y)}^{f_{X,Y}(x,y)} dy \quad \text{Thm.}$$

$$\mathbf{E}[X | Y = y] = \sum_x x p_{X|Y}(x|y)$$

$$\mathbf{E}[X | Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \quad \text{Def.}$$

$$\mathbf{E}[X] = \sum_y p_Y(y) \mathbf{E}[X | Y = y]$$

$$\begin{aligned} \mathbf{E}[X] &= \int_{-\infty}^{\infty} f_Y(y) \mathbf{E}[X | Y = y] dy \\ &= \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx dy \end{aligned}$$

- Expected value rule...

$$= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} \cancel{f_Y(y)} f_{X|Y}(x|y) dy dx$$

$$= \int_{-\infty}^{\infty} x f_X(x) dx = \mathbf{E}[x]$$

$$\mathbf{E}[g(x) | Y = y]$$

$$= \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx$$

## Independence

$$p_{X,Y}(x,y) = p_X(x) p_Y(y), \quad \text{for all } x, y$$

$$f_{X,Y}(x,y) = \underline{f_X(x)} f_Y(y), \quad \text{for all } x \text{ and } y$$

$$f_{Y|X} = f_Y$$

$$f_{X,Y}(x,y) = \underline{f_{X|Y}(x|y)} f_Y(y)$$

- equivalent to:  $f_{X|Y}(x|y) = f_X(x)$ , for all  $y$  with  $f_Y(y) > 0$  and all  $x$

If  $X, Y$  are **independent**:  $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

$$\mathbf{var}(X + Y) = \mathbf{var}(X) + \mathbf{var}(Y)$$

$g(X)$  and  $h(Y)$  are also independent:  $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

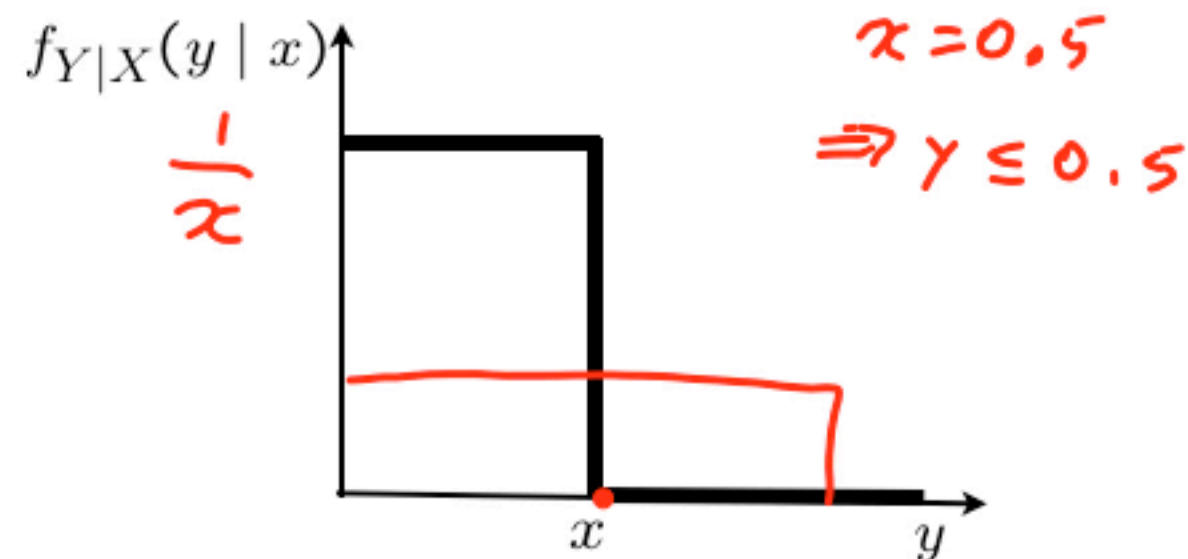
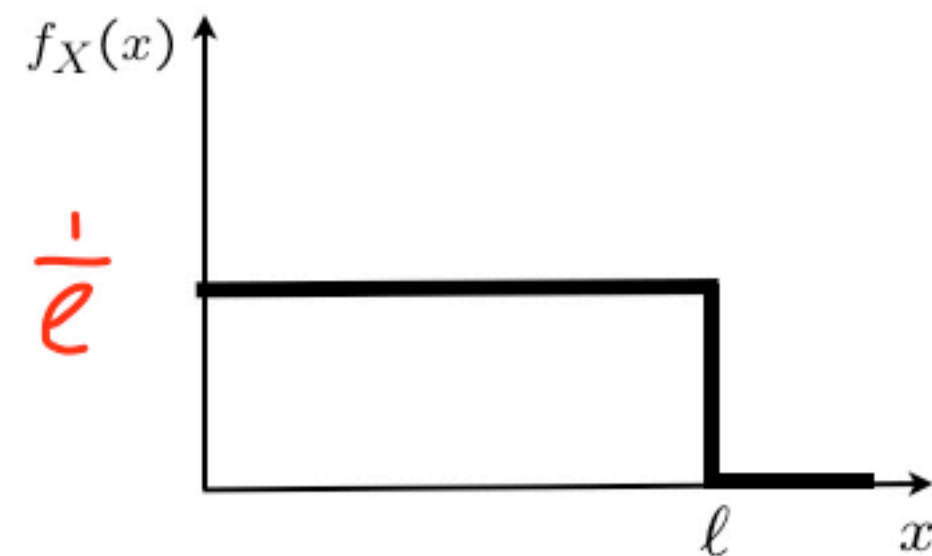
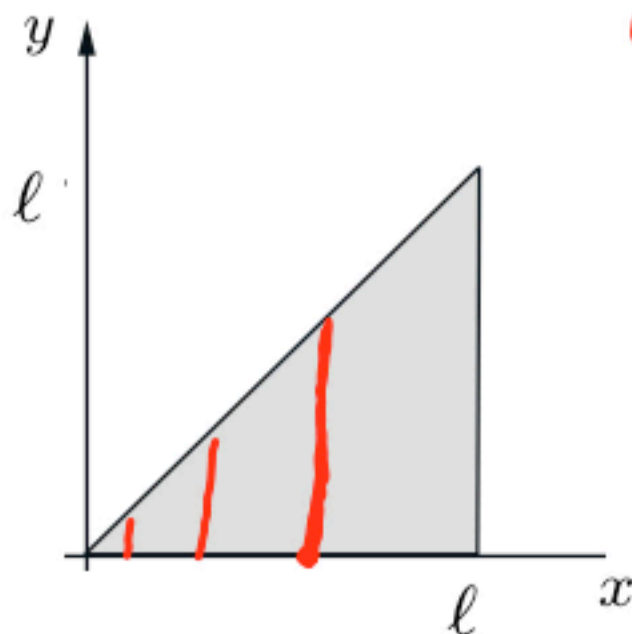
## Stick-breaking example



- Break a stick of length  $\ell$  twice
  - first break at  $X$ : uniform in  $[0, \ell]$
  - second break at  $Y$ : uniform in  $[0, X]$

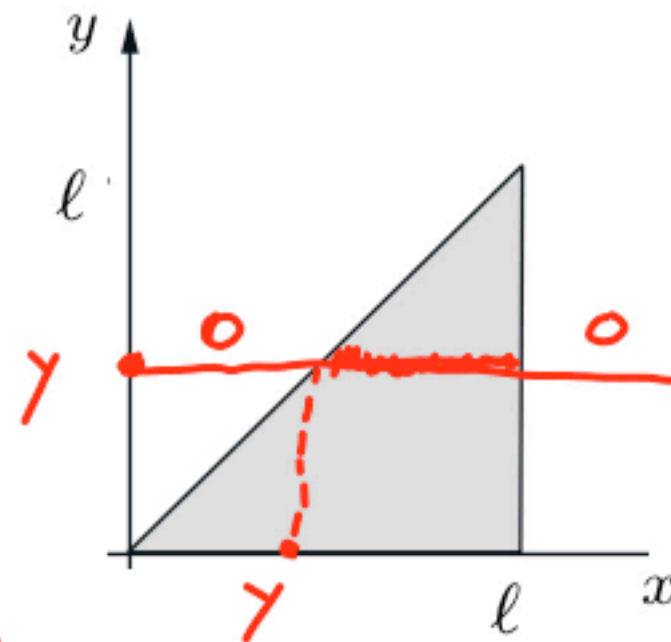
$$f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x) = \frac{1}{\ell x}$$

$$0 \leq y \leq x \leq \ell$$



## Stick-breaking example

$$f_{X,Y}(x,y) = \frac{1}{\ell x}, \quad 0 \leq y \leq x \leq \ell$$

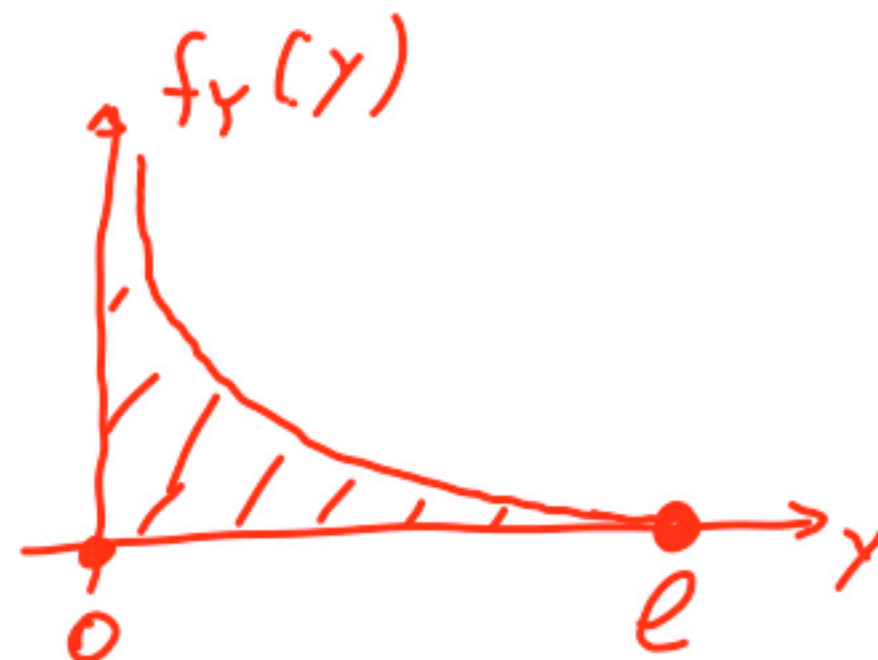


$$f_Y(y) = \int_y^\ell f_{X,Y}(x,y) dx = \int_y^\ell \frac{1}{\ell x} dx = \frac{1}{\ell} \log\left(\frac{\ell}{y}\right)$$

$$\mathbf{E}[Y] = \int_0^\ell y \frac{1}{\ell} \log\left(\frac{\ell}{y}\right) dy$$

- Using total expectation theorem:

$$\mathbf{E}[Y] = \int_0^\ell \frac{1}{\ell} \mathbf{E}[Y|X=x] dx = \int_0^\ell \left(\frac{1}{\ell}\right) \frac{x}{2} dx = \frac{1}{2} \mathbf{E}[X] = \frac{1}{2} \cdot \frac{\ell}{2} = \frac{\ell}{4}.$$



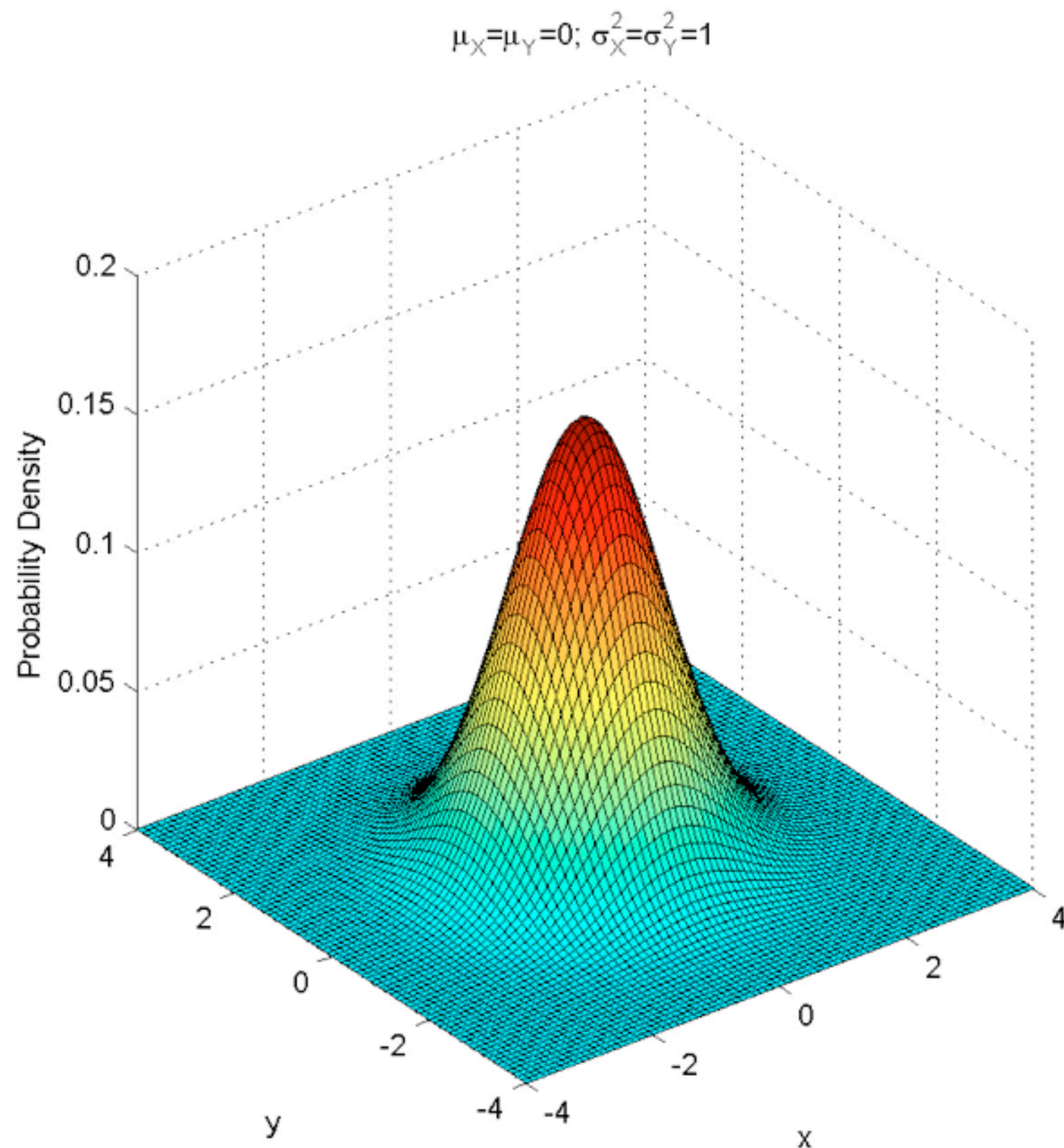
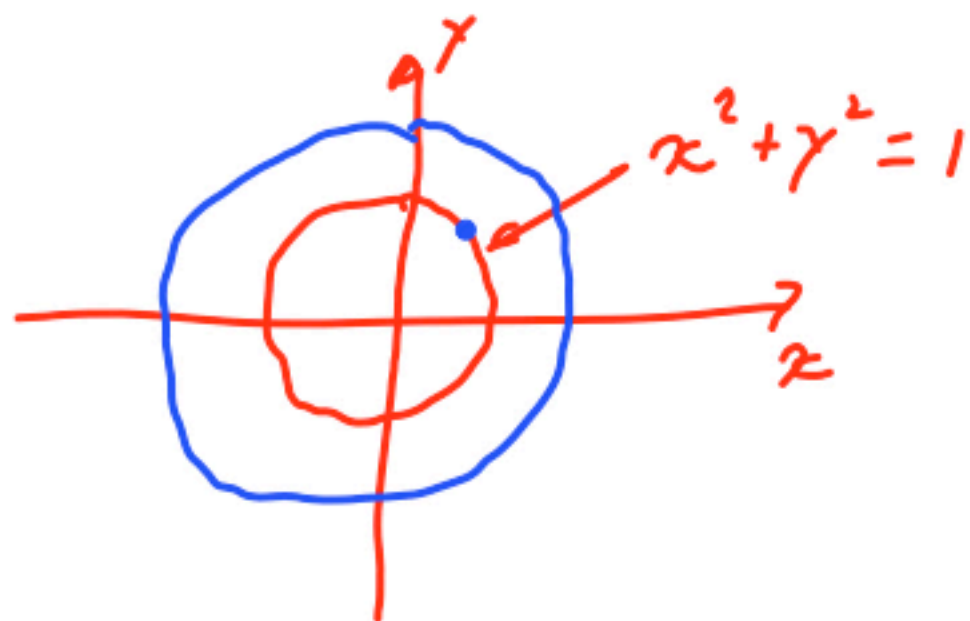


## Independent standard normals

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\}$$

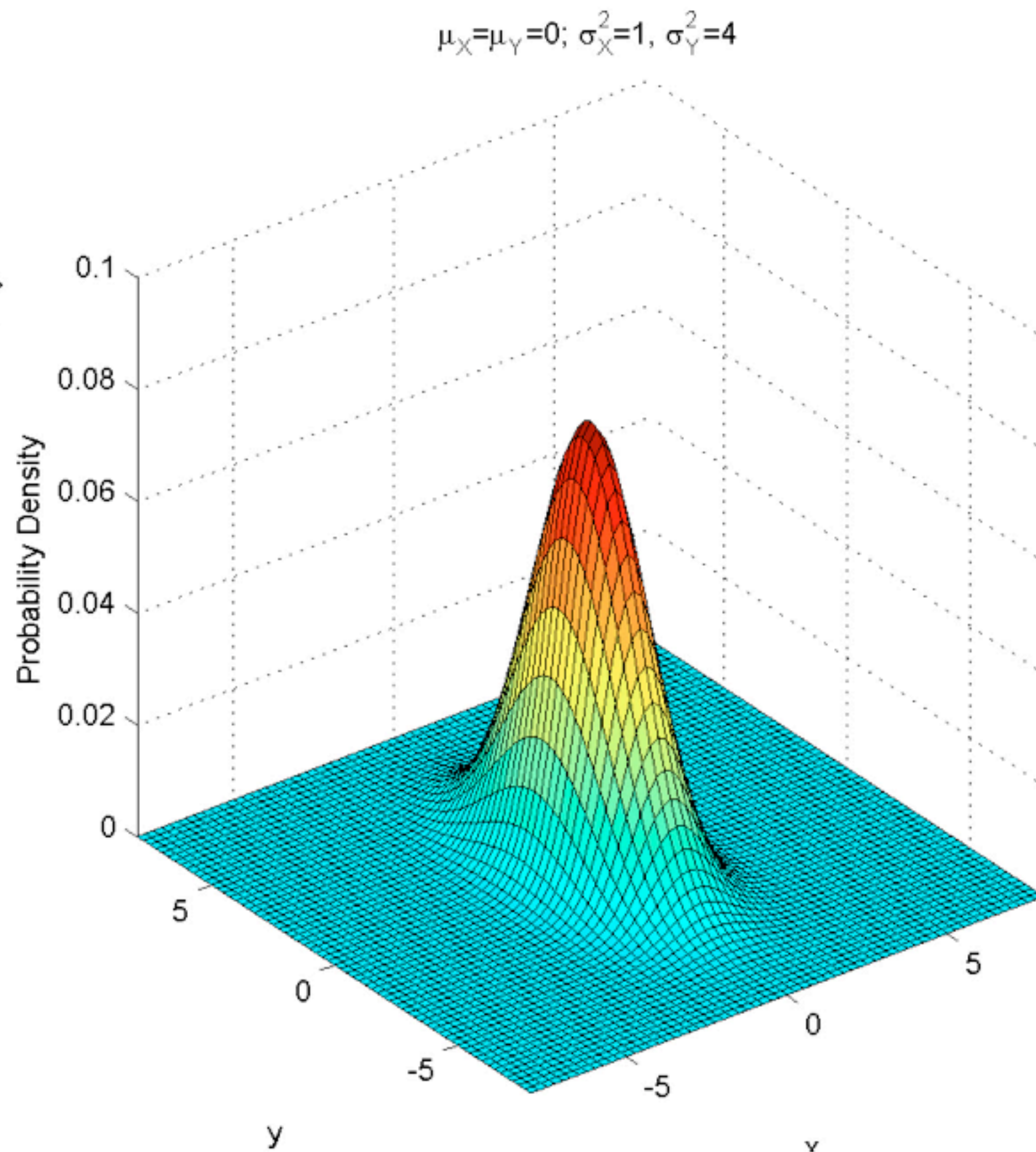
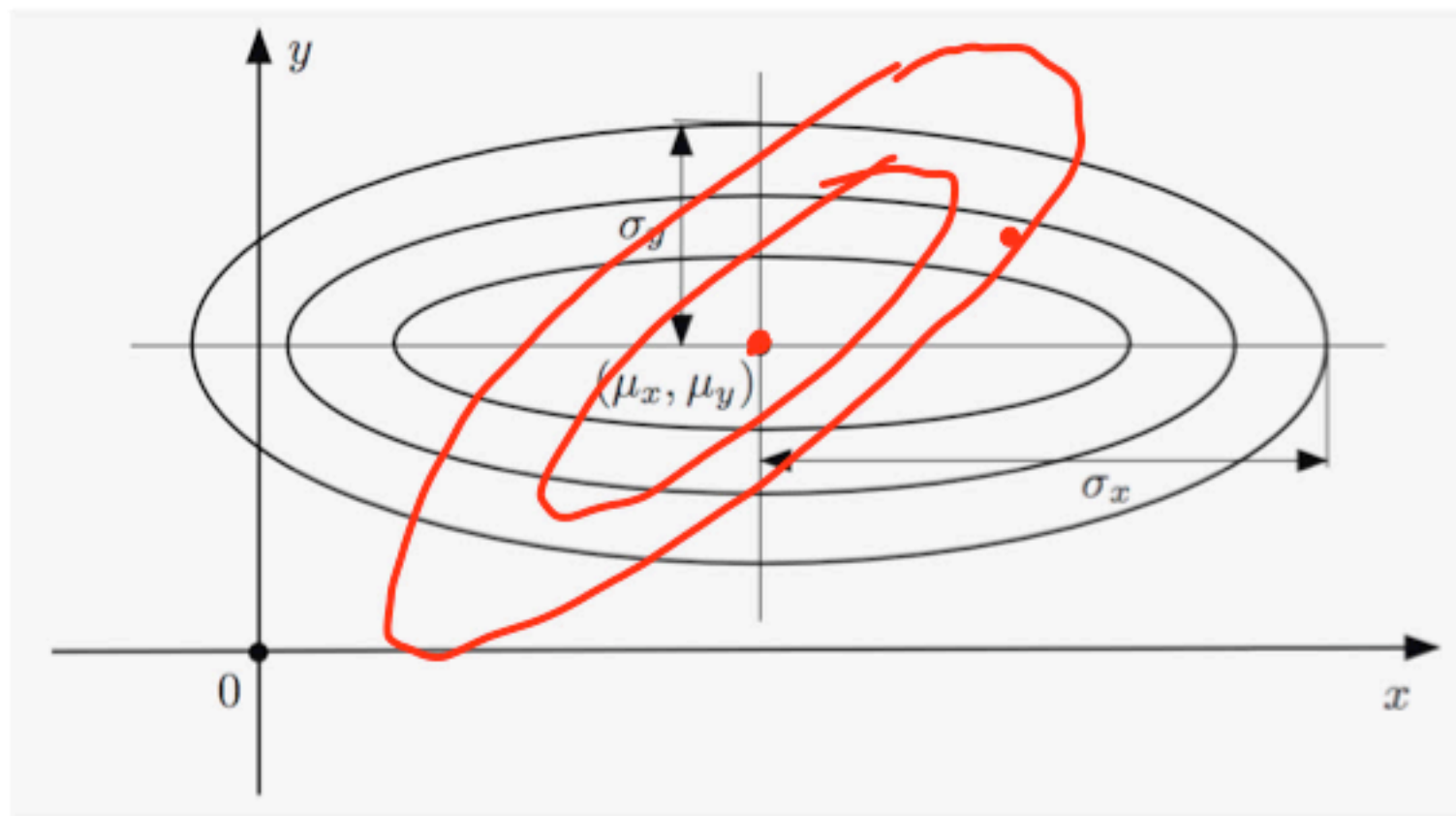
$$= \frac{1}{2\pi} \exp\left\{-\frac{1}{2}(x^2 + y^2)\right\}$$



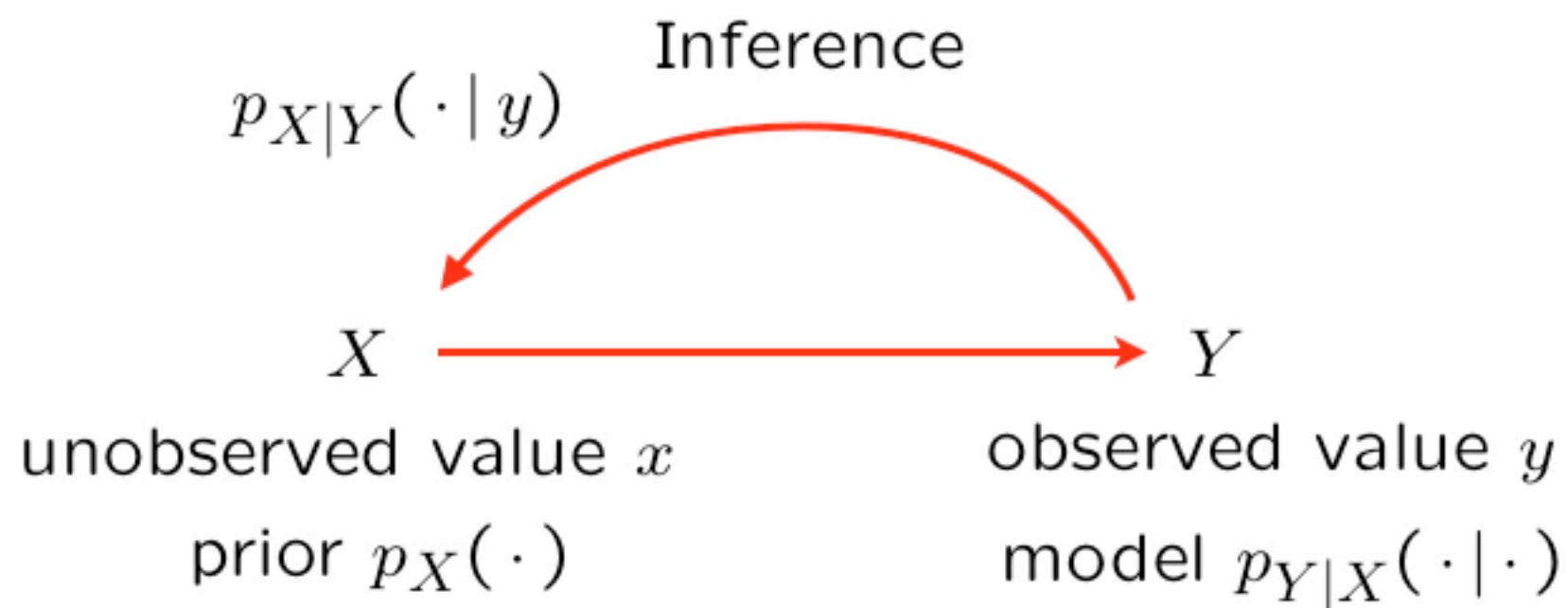


## Independent normals

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$
$$= \underbrace{\frac{1}{2\pi\sigma_x\sigma_y}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right\}$$



## The Bayes rule — a theme with variations



$$\begin{aligned} p_{X,Y}(x, y) &= p_X(x) p_{Y|X}(y | x) \\ &= p_Y(y) p_{X|Y}(x | y) \end{aligned}$$

$$\begin{aligned} f_{X,Y}(x, y) &= f_X(x) f_{Y|X}(y | x) \\ &= f_Y(y) f_{X|Y}(x | y) \end{aligned}$$

$$p_{X|Y}(x | y) = \frac{p_X(x) p_{Y|X}(y | x)}{p_Y(y)}$$

$$f_{X|Y}(x | y) = \frac{f_X(x) f_{Y|X}(y | x)}{f_Y(y)}$$

*posterior*

$$p_Y(y) = \sum_{x'} p_X(x') p_{Y|X}(y | x')$$

$$f_Y(y) = \int f_X(x') f_{Y|X}(y | x') dx' \bullet$$

## The Bayes rule — one discrete and one continuous random variable

$K$ : discrete

$Y$ : continuous

$$P(K=k, \gamma \leq Y \leq \gamma + \delta) \quad \delta > 0, \delta \approx 0$$

$$= P(K=k) P(\gamma \leq Y \leq \gamma + \delta | K=k) \approx p_K(k) f_{Y|K}(\gamma | k) \delta$$

$$= P(\gamma \leq Y \leq \gamma + \delta) P(K=k | \gamma \leq Y \leq \gamma + \delta) \approx f_Y(\gamma) \delta p_{K|Y}(k | \gamma)$$

$$p_{K|Y}(k | y) = \frac{p_K(k) f_{Y|K}(y | k)}{f_Y(y)}$$

$$f_{Y|K}(y | k) = \frac{f_Y(y) p_{K|Y}(k | y)}{p_K(k)}$$

$$f_Y(y) = \sum_{k'} p_K(k') f_{Y|K}(y | k')$$

$$p_K(k) = \int f_Y(y') p_{K|Y}(k | y') dy'$$



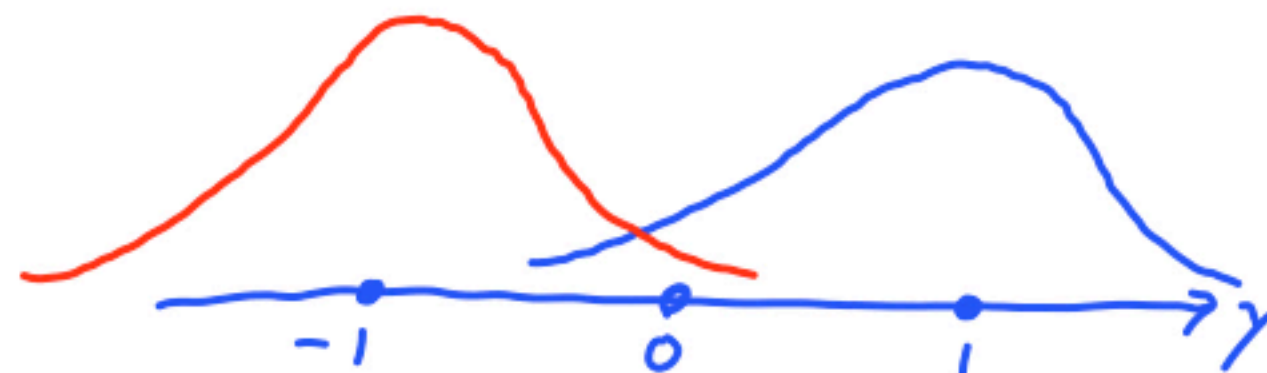
## The Bayes rule — discrete unknown, continuous measurement

- unknown  $K$ : equally likely to be  $-1$  or  $+1$
- measurement  $Y$ :  $Y = K + W$ ;  $W \sim \mathcal{N}(0, 1)$



$$Y|K=1 \sim \mathcal{N}(1, 1)$$

$$Y|K=-1 \sim \mathcal{N}(-1, 1)$$



- Probability that  $K = 1$ , given that  $Y = y$ ?

$$p_{K|Y}(1|y)$$

$$p_K(k) = 1/2 \quad f_{Y|K}(y|k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-k)^2}$$

$k = -1, +1$

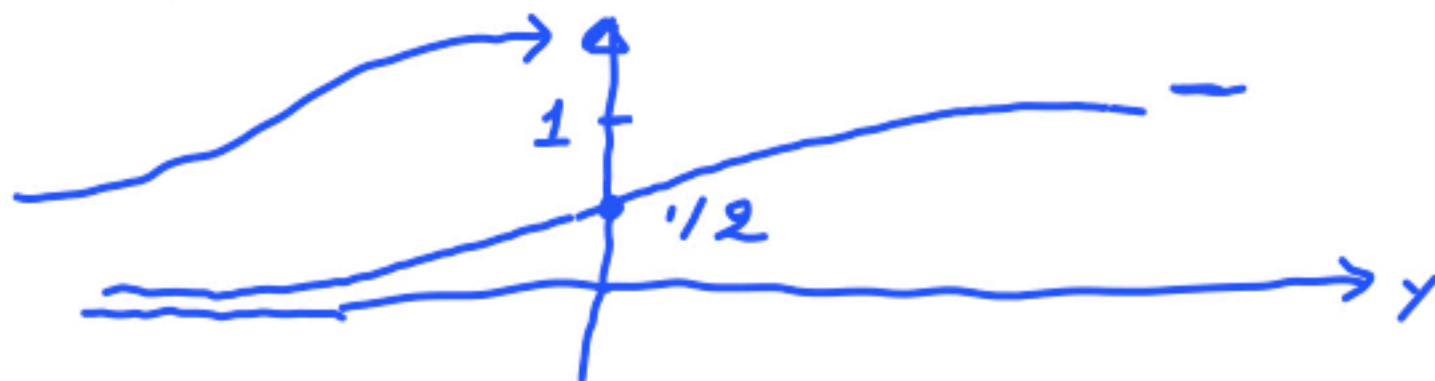
$$p_{K|Y}(k|y) = \frac{p_K(k) f_{Y|K}(y|k)}{f_Y(y)}$$

$$f_Y(y) = \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y+1)^2} + \frac{1}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-1)^2}$$

$$f_Y(y) = \sum_{k'} p_K(k') f_{Y|K}(y|k')$$

$$p_{K|Y}(1|y) = \frac{1}{1 + e^{-2y}}$$

algebra



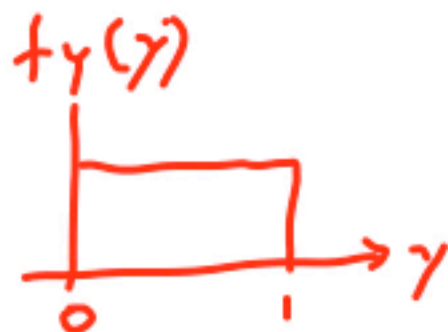
## The Bayes rule — continuous unknown, discrete measurement

- measurement  $K$ : Bernoulli with parameter  $Y$



$$f_{Y|K}(y | \underset{1}{k}) = \frac{f_Y(y) p_{K|Y}(k | y)}{p_K(k)}$$

- unknown  $Y$ : uniform on  $[0, 1]$



$$p_K(k) = \int f_Y(y') p_{K|Y}(k | y') dy'$$

- Distribution of  $Y$  given that  $K = 1$ ?

$$f_{Y|K}(y | 1)$$

$$f_Y(y) = \begin{cases} 1 & y \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

$$p_{K|Y}(1 | y) = y$$

$$p_K(1) = \int_0^1 1 \cdot y dy = \left. \frac{y^2}{2} \right|_0^1 = \frac{1}{2}$$

$$f_{Y|K}(y | 1) = \frac{1 \cdot y}{1/2} = 2y, \quad y \in [0, 1]$$

