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.

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

$$\begin{array}{ll} p_{X,Y}(x,y) & f_{X,Y}(x,y) \\ p_{X|A}(x) & f_{X|A}(x) \\ p_{X|Y}(x \mid y) & f_{X|Y}(x \mid y) \end{array}$$

Definition: $\quad f_{X|Y}(x \mid y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{if } f_Y(y) > 0$

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

$$Y = y \qquad Y \approx y$$

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

Definition: $\quad \mathbf{P}(X \in A \mid Y = y) = \displaystyle\int_A f_{X|Y}(x \mid y)\,dx$
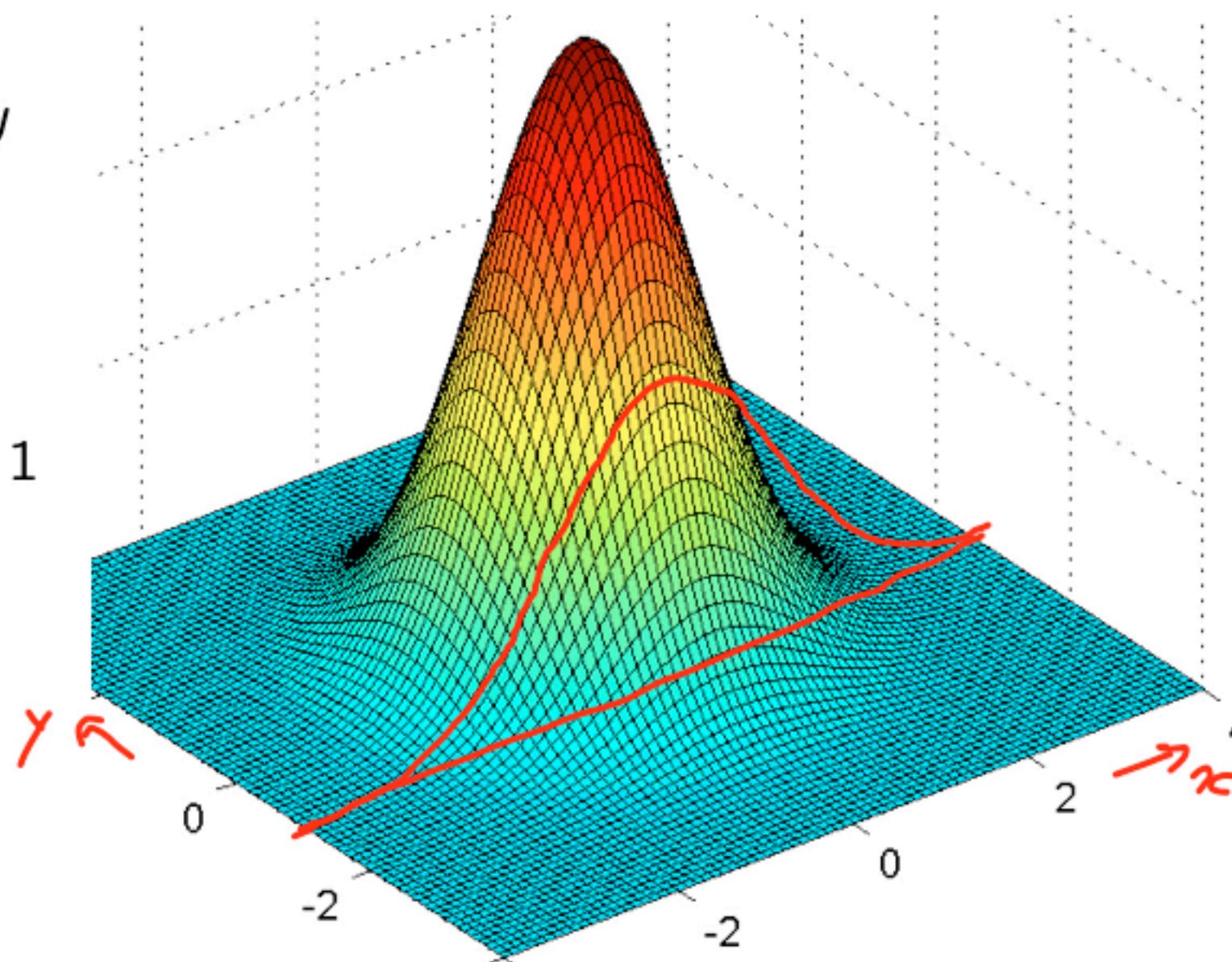
Comments on conditional PDFs

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

- $f_{X|Y}(x \mid y) \geq 0$

- Think of value of Y as fixed at some y

  shape of $f_{X|Y}(\cdot \mid y)$: slice of the joint

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

- Multiplication rule:

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

  $$= f_X(x) \cdot f_{Y|X}(y \mid x)$$

Total probability and expectation theorems

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

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

$$\overbrace{\phantom{f_Y(y)f_{X|Y}(x \mid y)}}^{f_{X,Y}(x,y)}$$

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

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

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

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} f_Y(y)\mathbf{E}[X \mid Y = y]\,dy$$

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

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

- Expected value rule…

$$E[g(X)|Y=y]$$

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

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

Independence

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

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

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

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

- 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:** $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

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

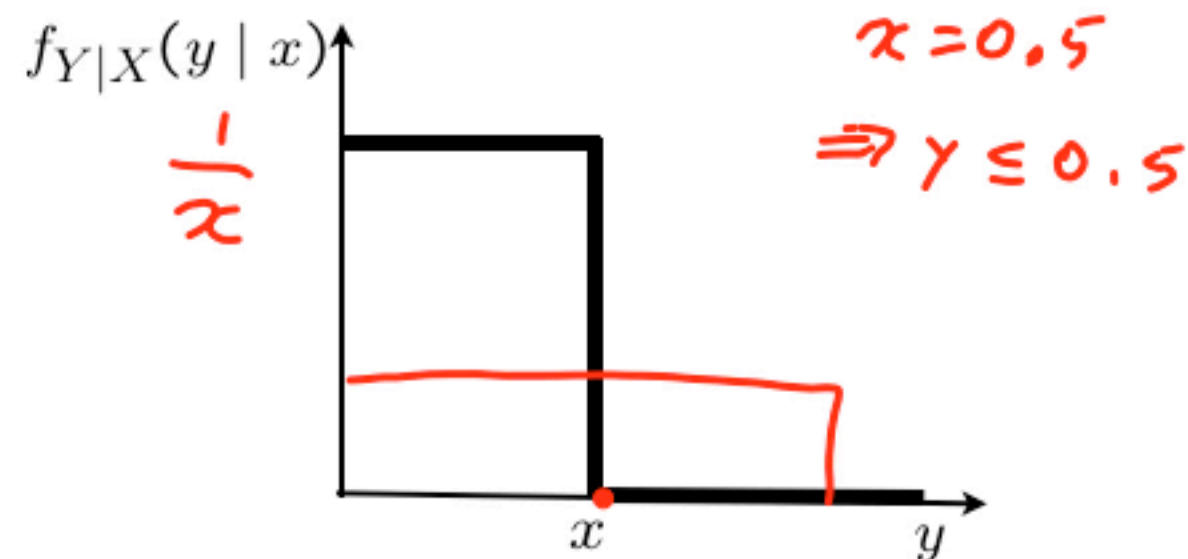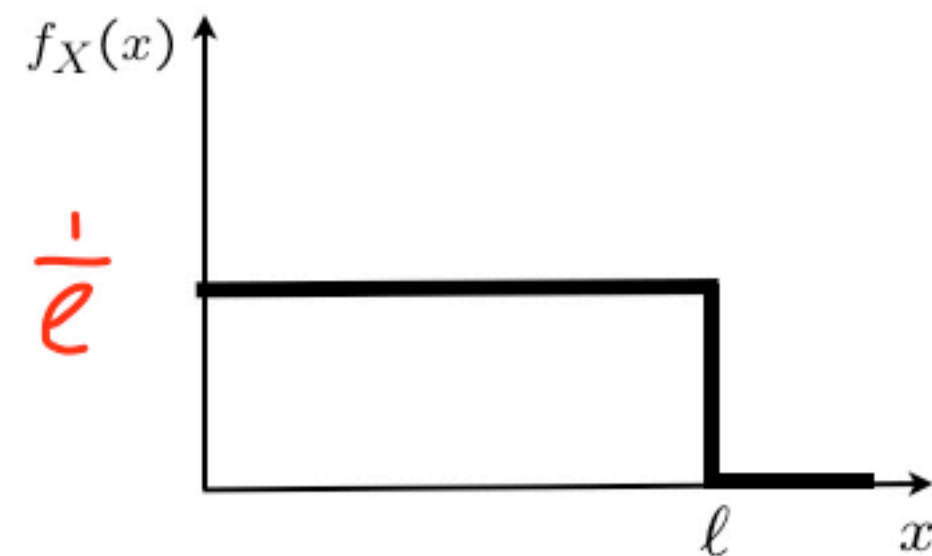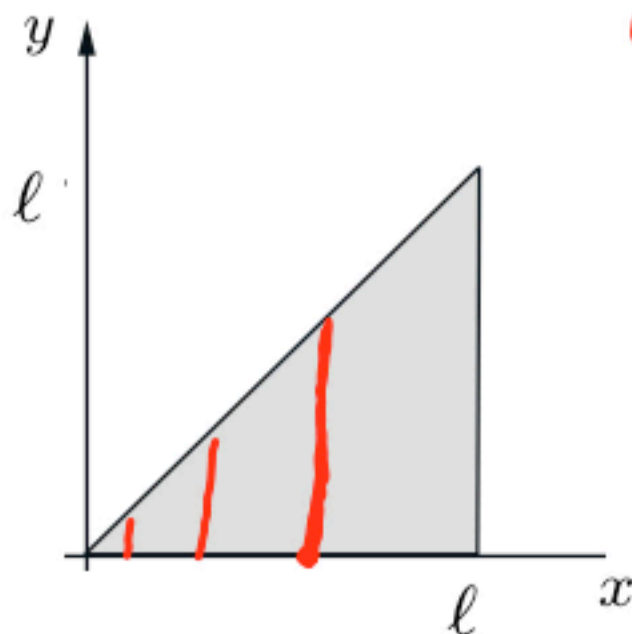$g(X)$ and $h(Y)$ are also independent: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

Stick-breaking example

0 r x ℓ

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

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

$$0 \le y \le x \le \ell$$

$f_X(x)$

$\dfrac{1}{\ell}$

ℓ x

$f_{Y|X}(y \mid x)$

$x = 0.5$

$\Rightarrow y \le 0.5$

$\dfrac{1}{x}$

x y

y

ℓ

ℓ x

Stick-breaking example

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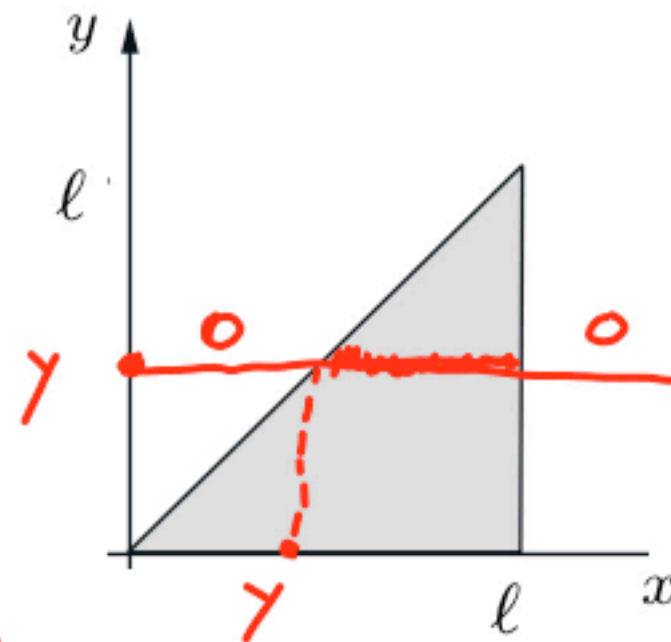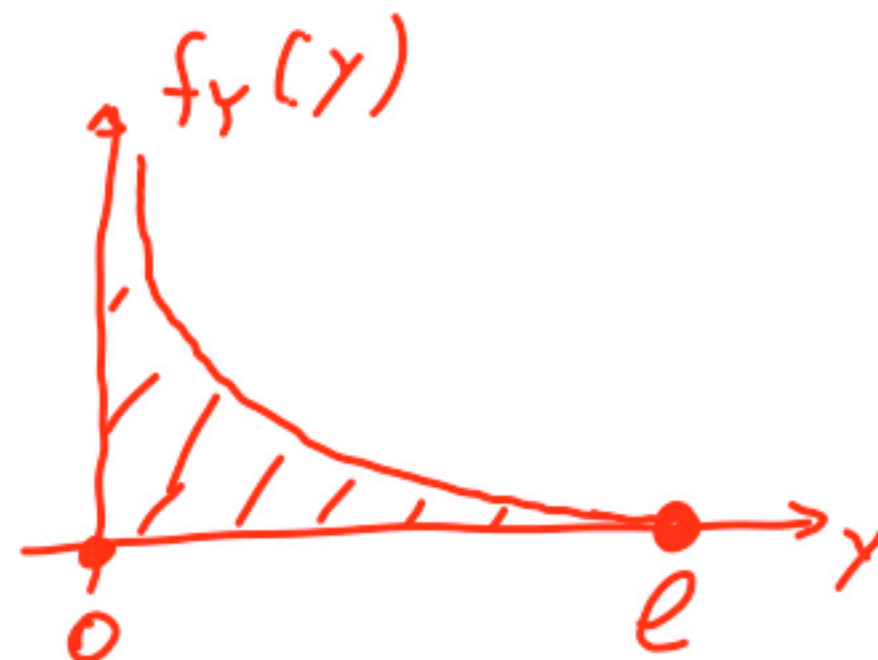

$$f_Y(y) = \int_y^\ell f_{X,Y}(x,y)\,dx = \int_y^\ell \frac{1}{\ell x}\,dx = \frac{1}{\ell}\log\left(\frac{\ell}{y}\right)$$

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

$f_Y(y)$

- Using total expectation theorem:

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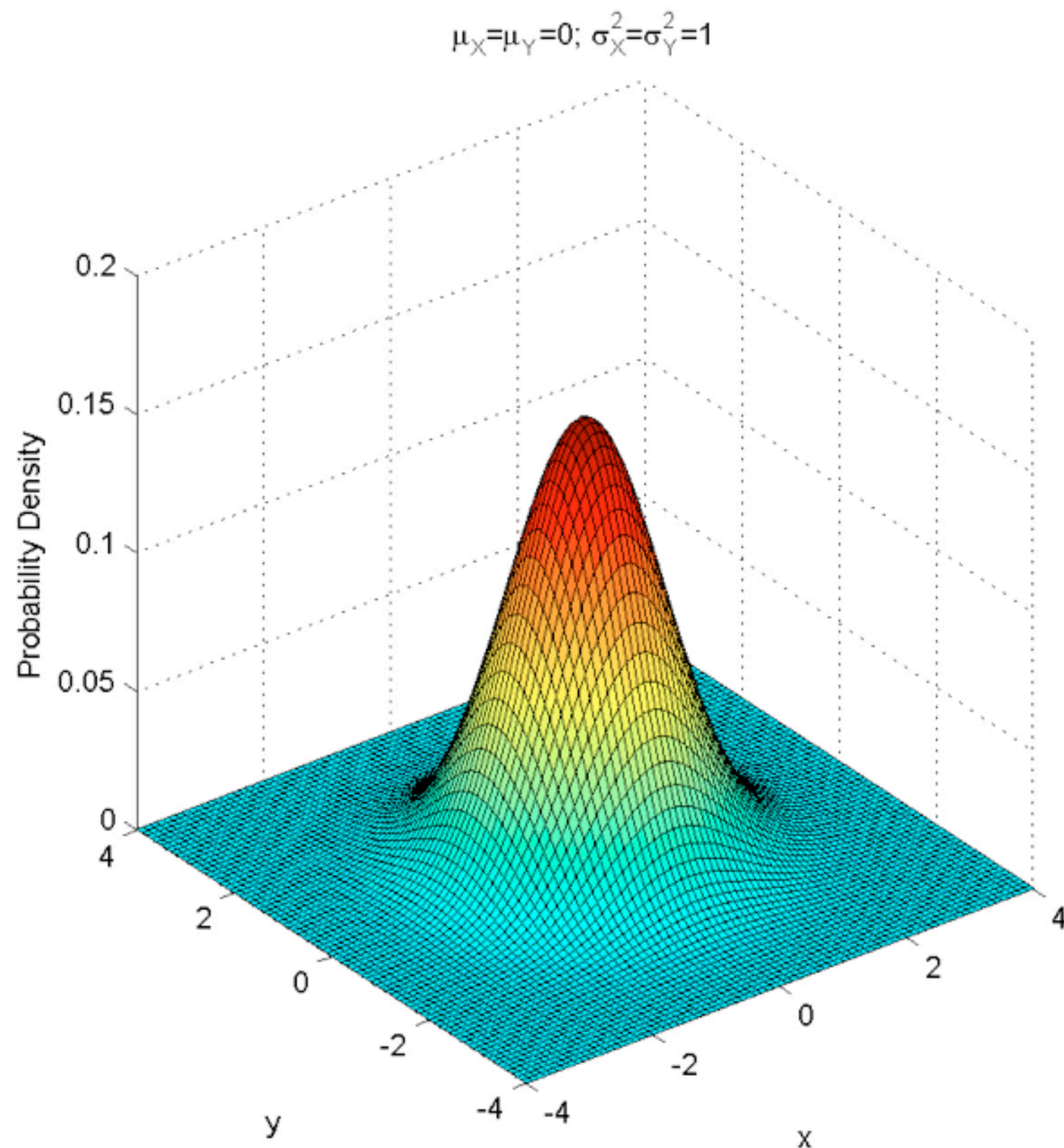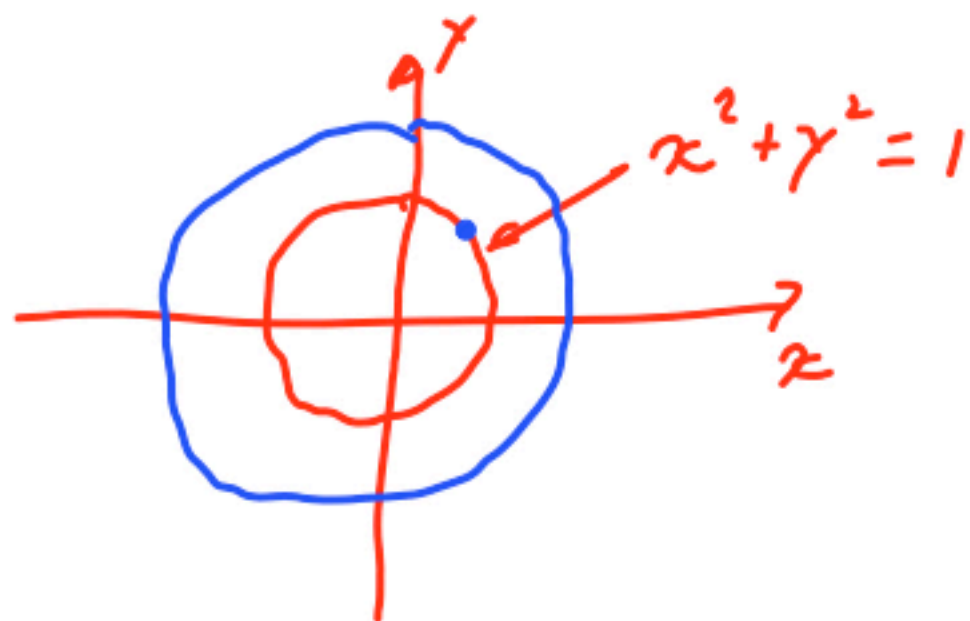
$f_x(x)$

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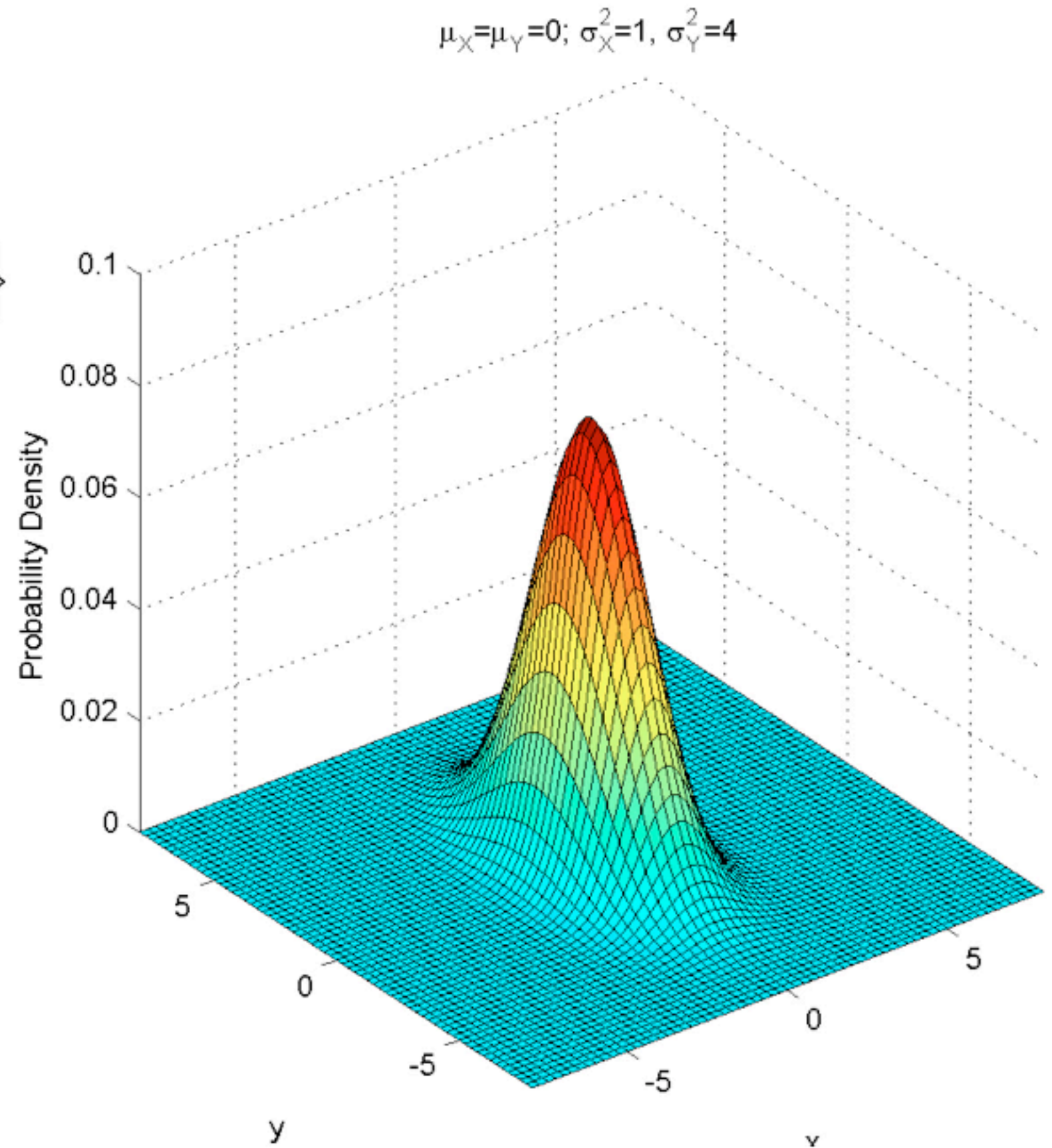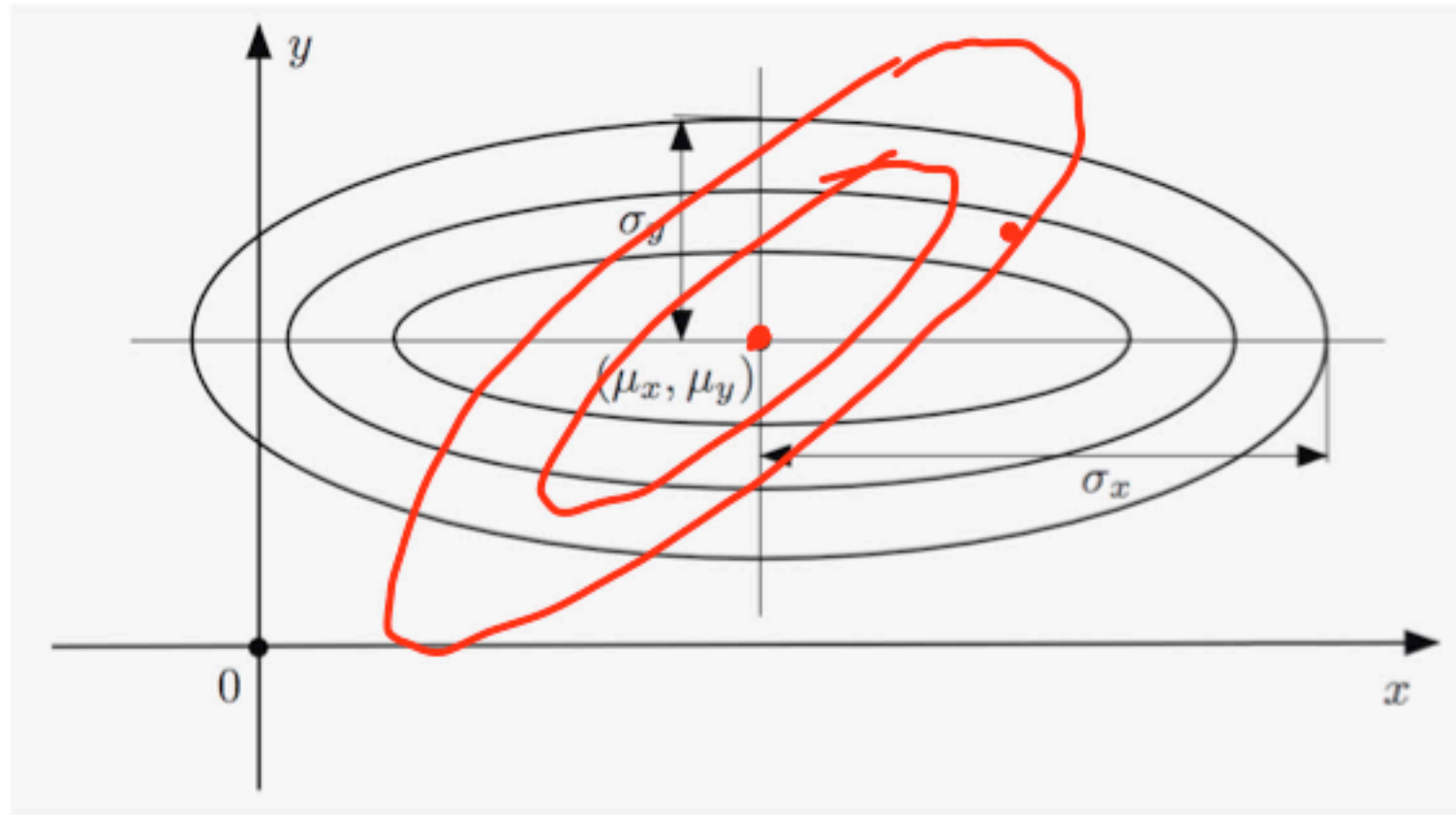
$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

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

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

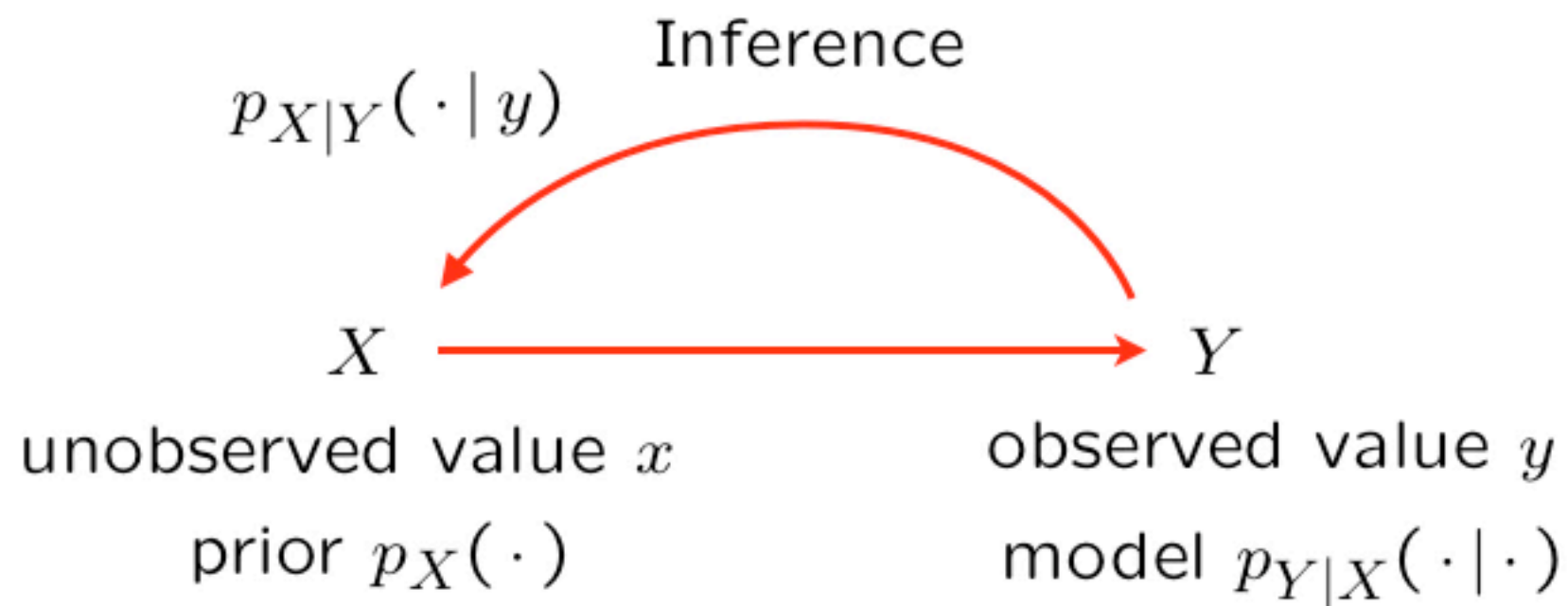$x^2 + y^2 = 1$

$\mu_X = \mu_Y = 0; \ \sigma_X^2 = \sigma_Y^2 = 1$

Independent normals

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

$$= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{ -\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2} \right\}$$

$\mu_X = \mu_Y = 0;\ \sigma_X^2 = 1,\ \sigma_Y^2 = 4$

The Bayes rule — a theme with variations

Inference

$p_{X|Y}(\cdot \,|\, y)$

$X \longrightarrow Y$

unobserved value x observed value y

prior $p_X(\cdot)$ model $p_{Y|X}(\cdot \,|\, \cdot)$

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

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

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

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

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

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

$$= P(K=k)\, P(y \leq Y \leq y+\delta \mid K=k) \quad \approx \quad p_K(k)\, f_{Y|K}(y|k)\, \delta$$

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

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

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

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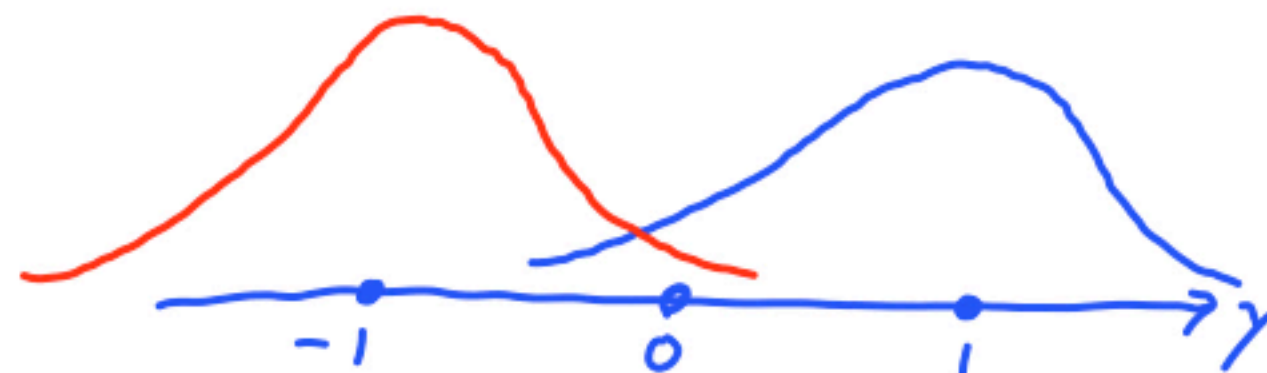
$$p_K(k) = \int f_Y(y')\, p_{K|Y}(k \mid y')\, dy'$$

The Bayes rule — discrete unknown, continuous measurement

- unkown K: equally likely to be -1 or $+1$

- measurement Y: $\quad Y = K + W; \quad W \sim \mathcal{N}(0,1)$

$K = \pm 1$

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

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

- Probability that $K = 1$, given that $Y = y$? $\quad p_{K|Y}(1|y)$

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

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

$k = -1, +1$

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

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

$$p_{K|Y}(1 \mid y) = \underset{algebra}{} \frac{1}{1 + e^{-2y}}$$
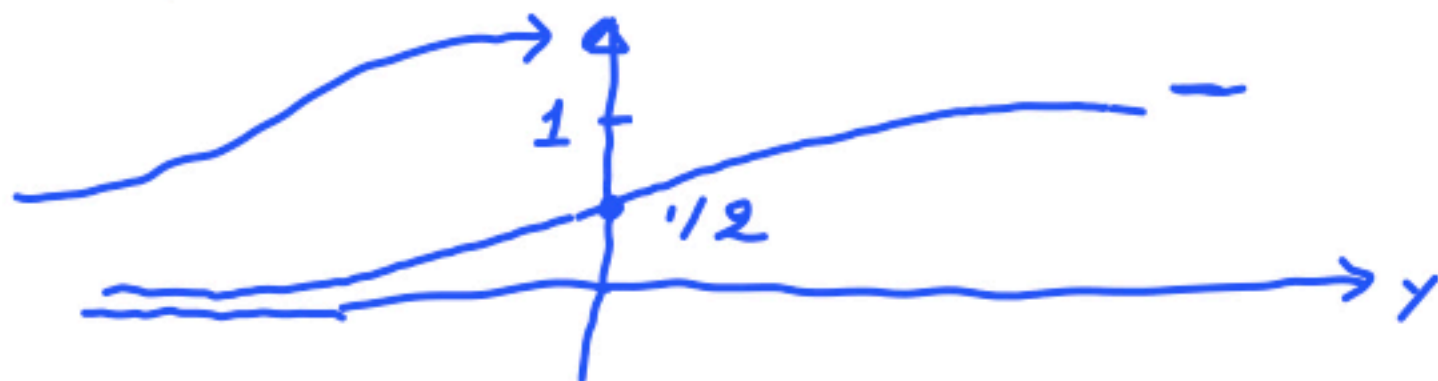
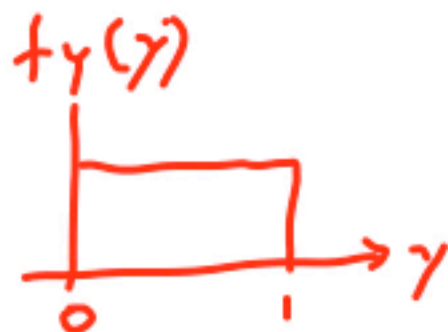# The Bayes rule — continuous unknown, discrete measurement

- measurement K: Bernoulli with parameter Y

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

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

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

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

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

$$f_Y(y) = 1 \quad y \in [0,1]$$
$$ 0 \quad \text{otherwise}$$

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

$$p_K(1) = \int_0^1 1 \cdot y \, dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2}$$

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