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.

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

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

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

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

$$\mathbf{P}(x \le X \le x + \delta \mid y \le Y \le y + \epsilon)$$

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

Comments on conditional PDFs

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

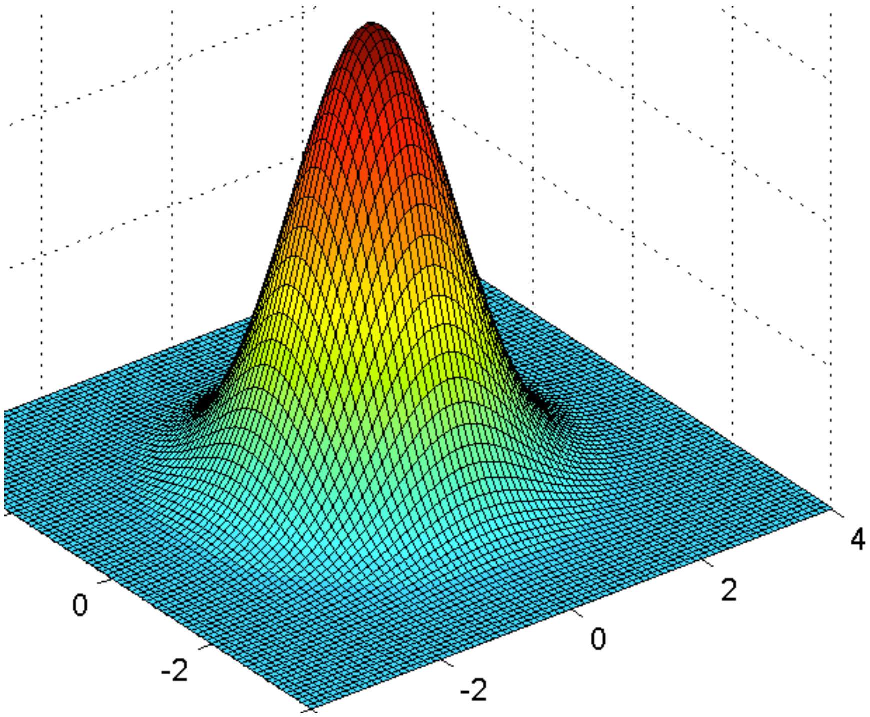
- $f_{X|Y}(x \mid y) \geq 0$

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

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

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

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



Total probability and expectation theorems

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x \,|\, y) \qquad f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x \,|\, y)\, dy$$

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

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

- Expected value rule...

Independence

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

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

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

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

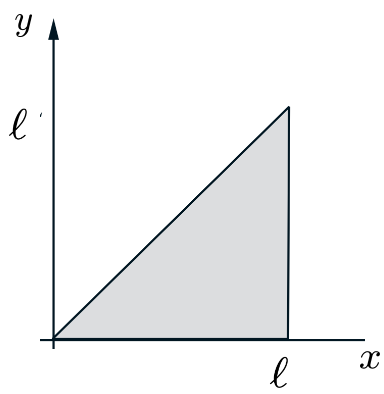
$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$

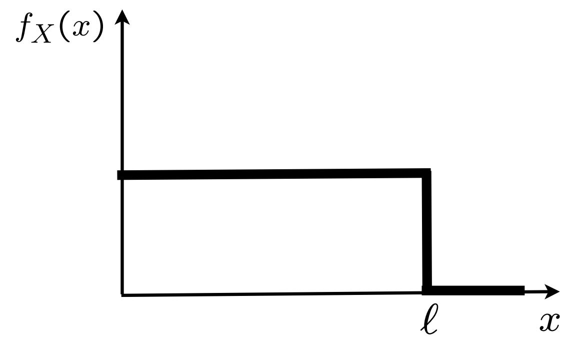
$g(X)$ and $h(Y)$ are also independent: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

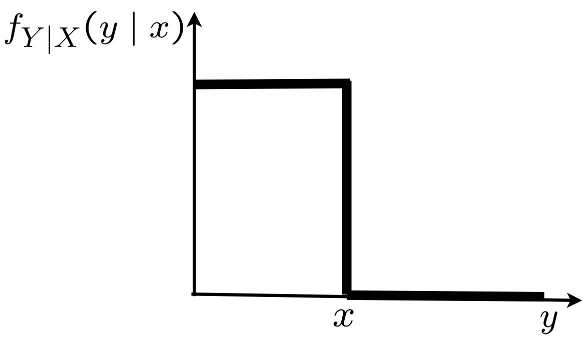
Stick-breaking example

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

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

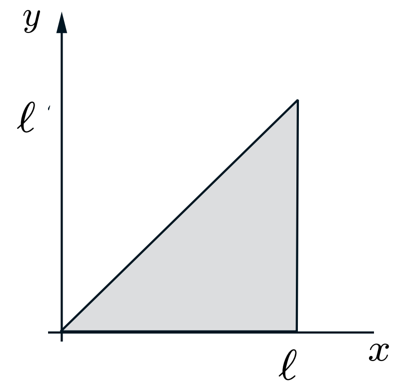






Stick-breaking example

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



$$f_Y(y) \ =$$

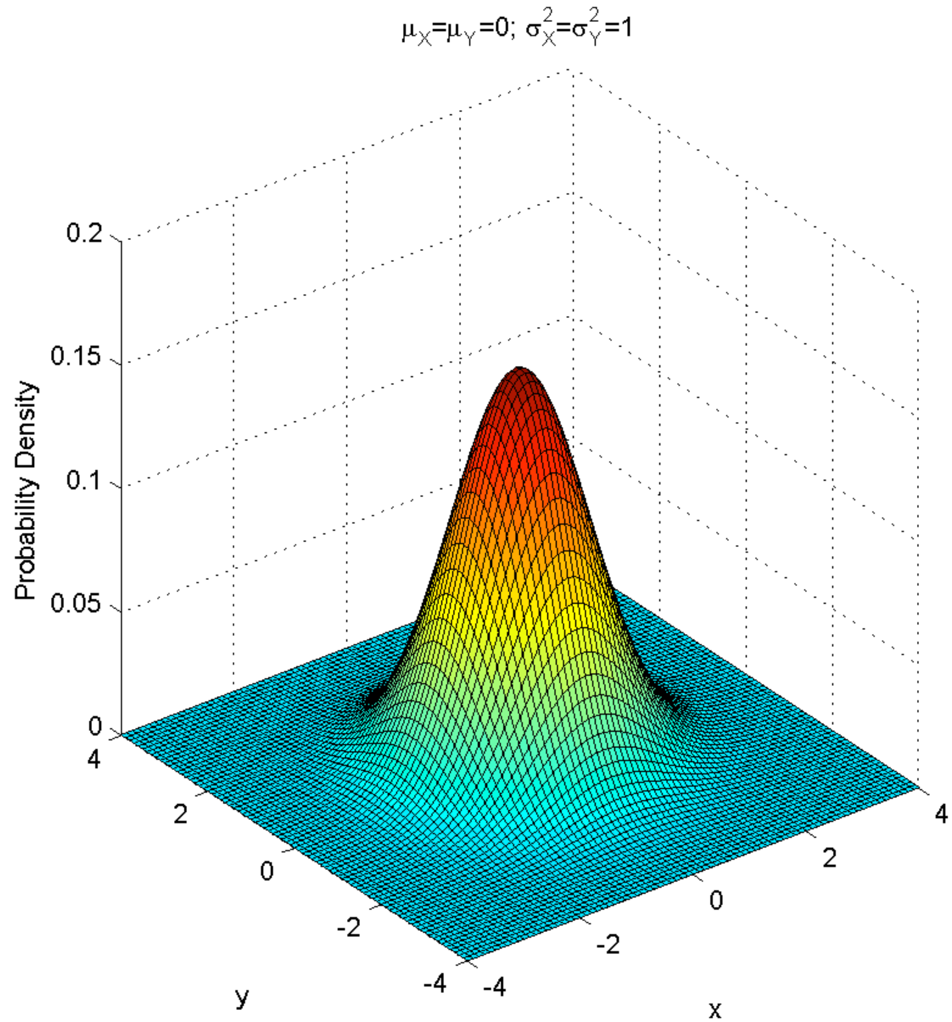
$$\mathbf{E}[Y] =$$

- Using total expectation theorem:

Independent standard normals

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

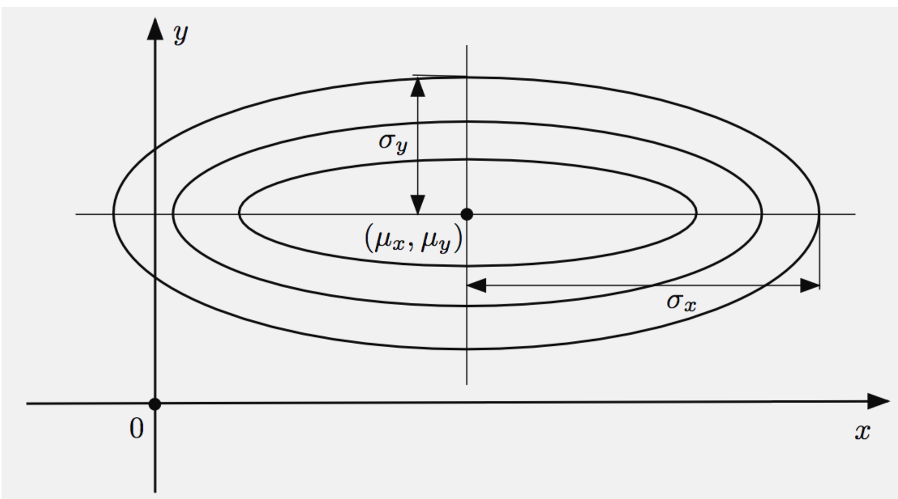
$$= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\}$$

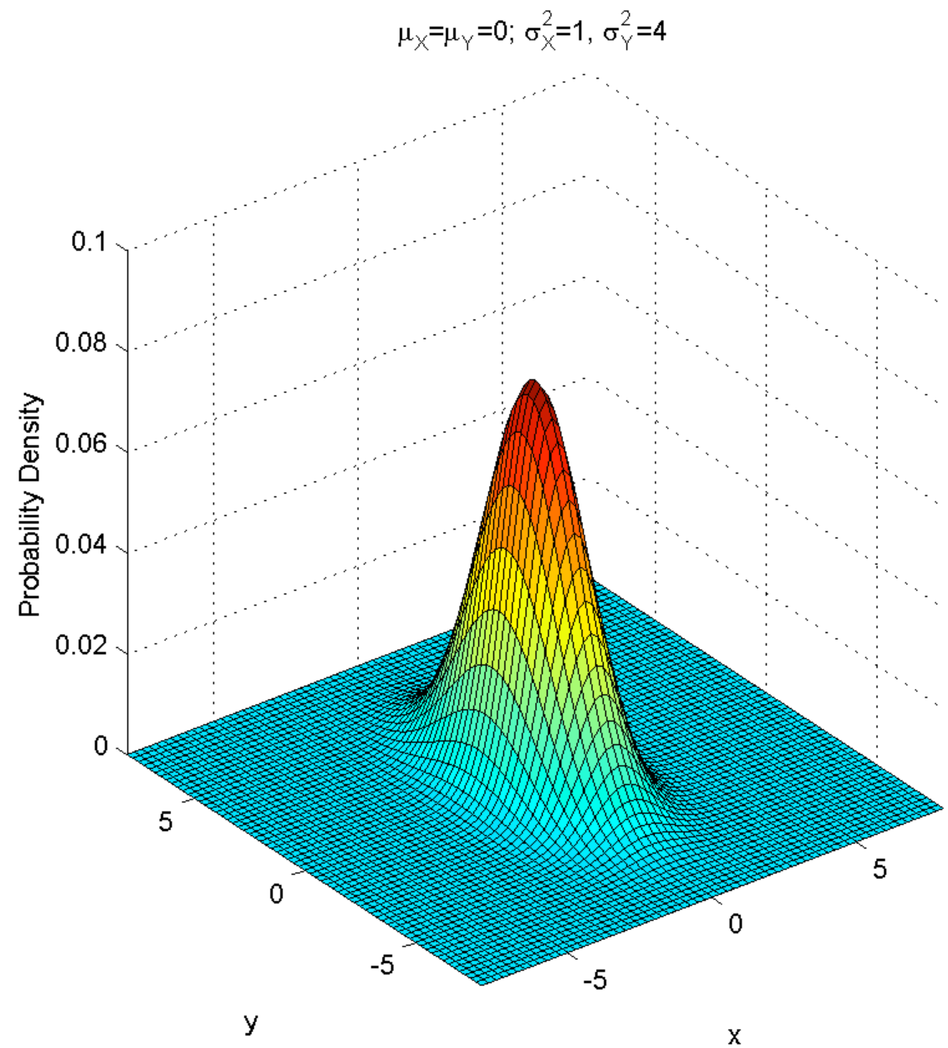


Independent normals

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

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





The Bayes rule — a theme with variations

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

$$\begin{tikzcd} X \arrow[r] & Y \arrow[l, bend right, "\text{Inference}"'] \end{tikzcd}$$

unobserved value x — observed value y

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

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

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

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

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

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

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

The Bayes rule — one discrete and one continuous random variable

K: discrete $\qquad$ Y: continuous

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

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

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

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

The Bayes rule — discrete unknown, continuous measurement

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

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

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

$p_K(k) =$ $\qquad$ $f_{Y|K}(y \,|\, k) =$

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

$f_Y(y) =$

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

The Bayes rule — continuous unknown, discrete measurement

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

- measurement K: Bernoulli with parameter Y

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

- unkown Y: uniform on $[0, 1]$
- Distribution of Y given that $K = 1$?

$$f_Y(y) =$$

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

$$p_K(1) =$$

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