ASEN 5044, Fall 2018

Statistical Estimation for Dynamical Systems

Lecture 12: Gaussian Distributions, Joint PDFs for Multiple Random Variables

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Tues 10/02/2018

University of Colorado Boulder

Announcements

- **HW 4 Due Thurs 10/4 at 11 am**
- Submit to Canvas
- **Some general notes from TA…**

$$\cancel{\begin{bmatrix} HF \\ HF^2 \\ \vdots \\ HF^n \end{bmatrix} \neq 0} \quad !! \qquad \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} = 0 \ \checkmark$$

- **Midterm 1: out this Thursday 10/4 [coverage: HWs 1-4]**
  - One week long take home exam posted to Canvas
  - Due Thurs 10/11/201~~7~~8 on Canvas by 11 am
  - Open book/notes – must complete by yourself (honor code applies)

- Prof. Ahmed out of town Tues afternoon 10/9 thru Thurs morning 10/11

Overview

- Last time: Expected values and Expectation operator

Today:

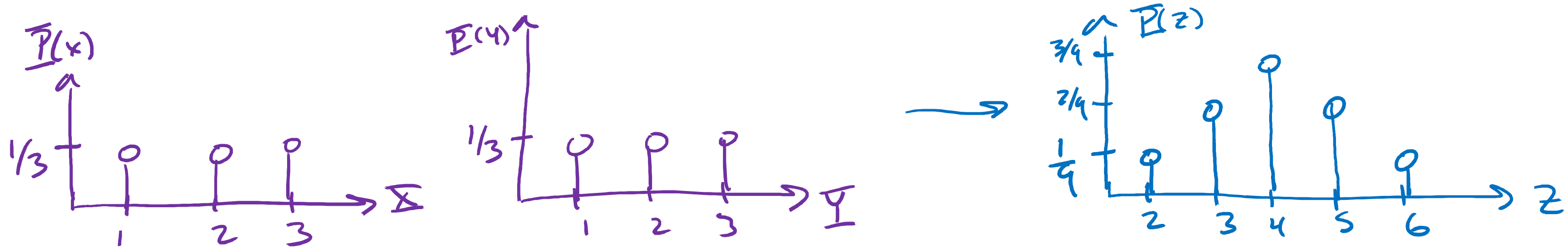
- Sums of independent random variables
  - Central limit theorem
- Gaussian (Normal) random variables and PDFs/distributions
- Joint PDFs for multiple random variables

  → Marginal and conditional pdfs for multiple random vars

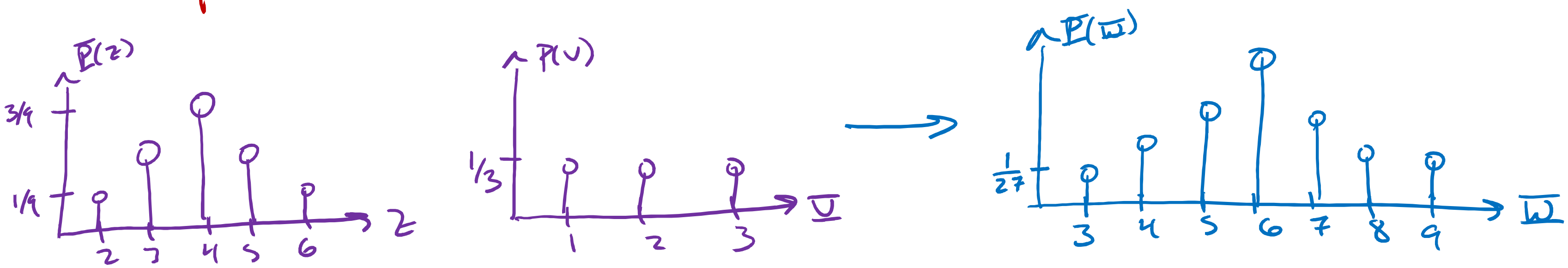
READ SIMON BOOK, CHAPTER 2.6

Sums of Independent Random Variables

- Suppose we have independent integer random variables X and Y with P(X) and P(Y)
- What's the distribution P(Z) of the sum Z = X + Y? (note: Z is another random variable!)



→ what if we added a 3rd RV $\perp\!\!\!\perp X$ & $\perp\!\!\!\perp Y$, e.g. $V \sim P(V)$? [So, get $W = X + Y + V = Z + V$]



→ Resulting discrete distributions for sum of 2 discrete RVs is a convolution of their dists!

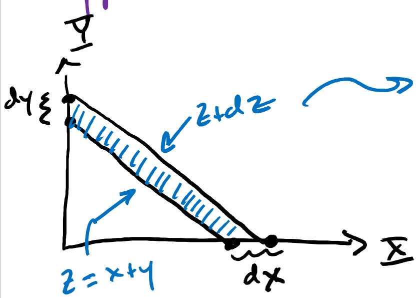
ie $P(Z) = P(X) * P(Y) = $ conv([$P_x(1), P_x(2), P_x(3)$], [$P_y(1), P_y(2), P_y(3)$])

↑ matlab

Sums of Independent Random Variables

- Frequently need to add *independent continuous* random variables X and Y
- If we know p(X) and p(Y), what does the pdf p(Z) for Z = X+Y look like?

Suppose $Z = z$ (some fixed #) : what are all the possible ways to get z from $X+Y$?
& " " " " " " " " $z+dz$ from $x+dx$ & $y+dy$



→ All x & y lying inside this differential strip map points b/w upper & lower bound lines to values b/w z & z+dz in z-space

$\Rightarrow P(z \le Z \le z+dz) = P(X \,\&\, Y$ lie inside differential strip$)$

$$= \iint_{\text{Diff. Strip}} p(x)\,p(y)\,dx\,dy \quad [\text{b/c } X \perp Y]$$

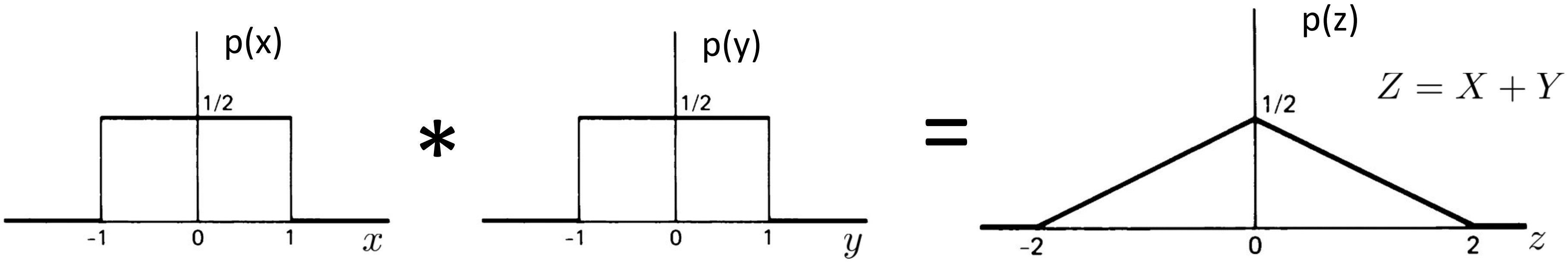
→ But we also know that since $x+y = z =$ some known # → $y = z - x$ → $dy = dz$ (for fixed x in integration)

→ $P(z \le Z \le z+dz) = \iint_{DS} p(x)\,p_y(z-x)\,dx\,dz$ → Integrate over all possible x & z: $P(\eta \le Z \le \xi) = \int_{\eta}^{\xi} \underbrace{\left[\int_{-\infty}^{\infty} p(x)\,p_y(z-x)\,dx\right]}_{p(z)} dz$

→ $\therefore\ p(z) = \int_{-\infty}^{\infty} p(x)\,p_y(z-x)\,dx = p_x(x) * p_y(y)$ = convolution of $p(x)$ & $p(y)$ [where $*$ = convolution integral op.]

Example: Sum of Two Uniform RVs

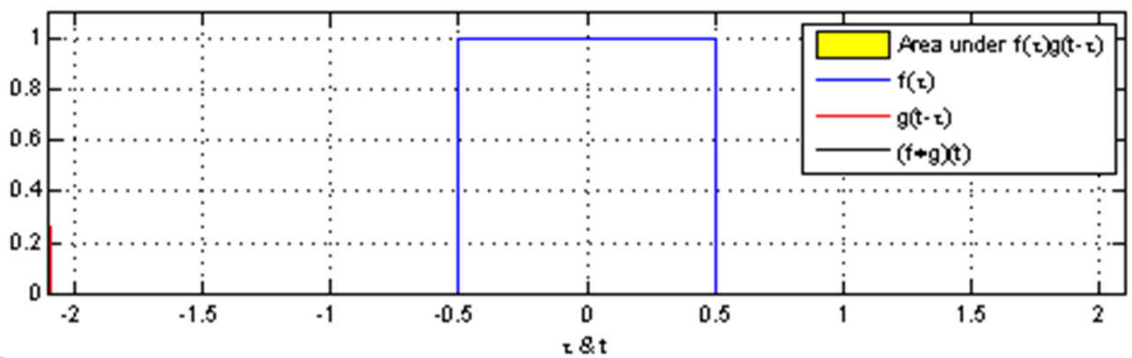
- Suppose p(X) = U(-1,1) and p(Y)=U(-1,1)



$$p(z) = \int_{-\infty}^{\infty} p_x(x)\, p_y(z-x)\, dx = \int_{-\infty}^{\infty} p_y(y)\, p_x(z-y)\, dy$$

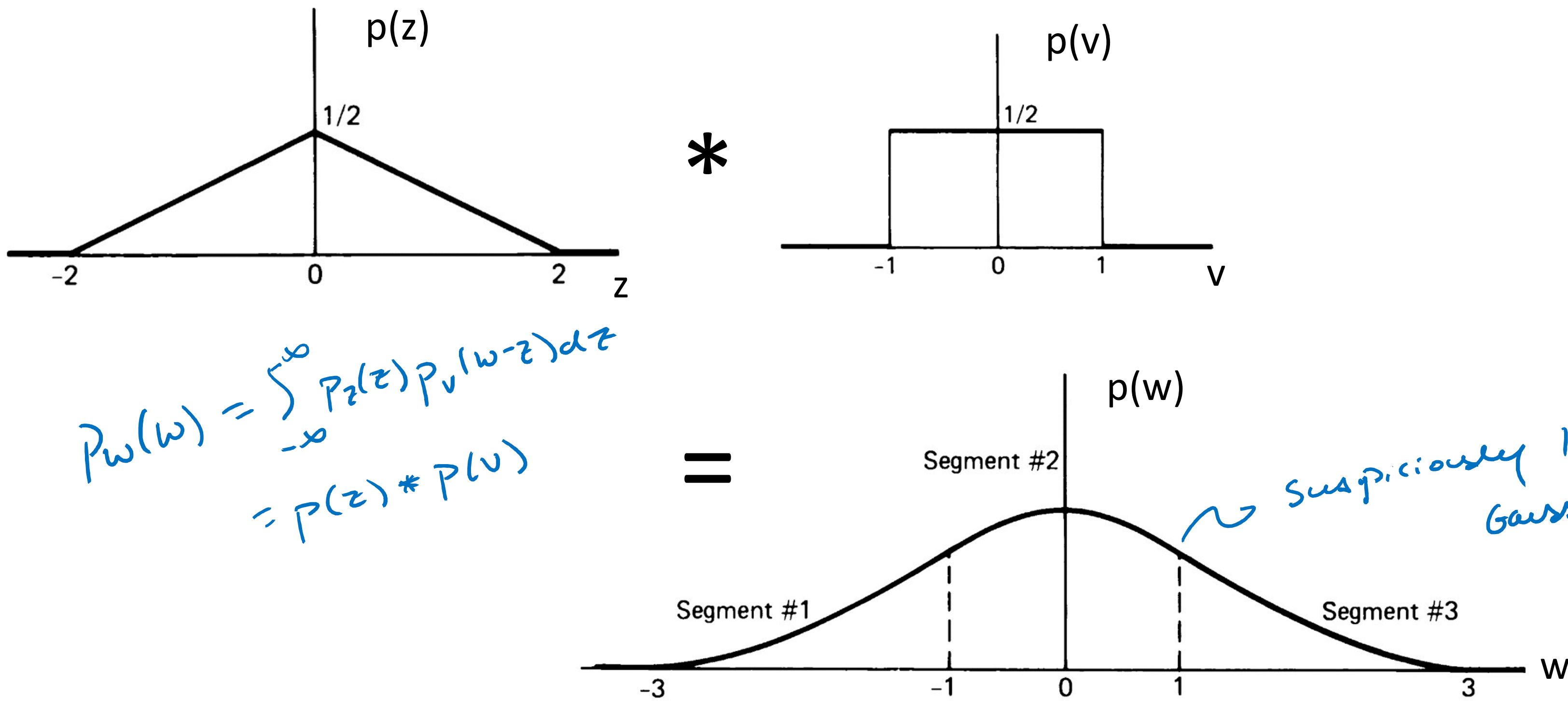
Animated Convolution Example (not to same scale as pdf figures)

https://commons.wikimedia.org/w/index.php?curid=11003835



Example: Sum of Three Uniform RVs

- Suppose X~U(-1,1), Y~U(-1,1), and V~U(-1,1), such that W = X + Y + V



$$p_w(w) = \int_{-\infty}^{\infty} p_z(z)\, p_v(w-z)\, dz$$
$$= p(z) * p(v)$$

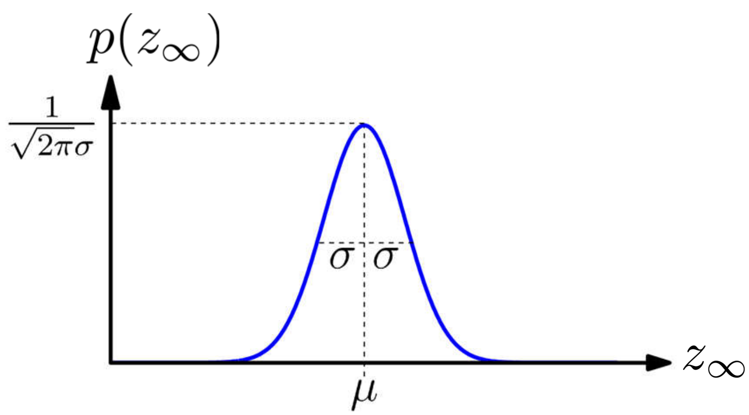
suspiciously looks Gaussian! . . .

The Central Limit Theorem

- If the sequence of RVs x_i , i=1,2,…,n,… consists of independent random variables, then (under some reasonably mild conditions) the pdf $p(z_n)$ of the sum

$$z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i$$

will tend to a **Gaussian pdf** as $n \to \infty$

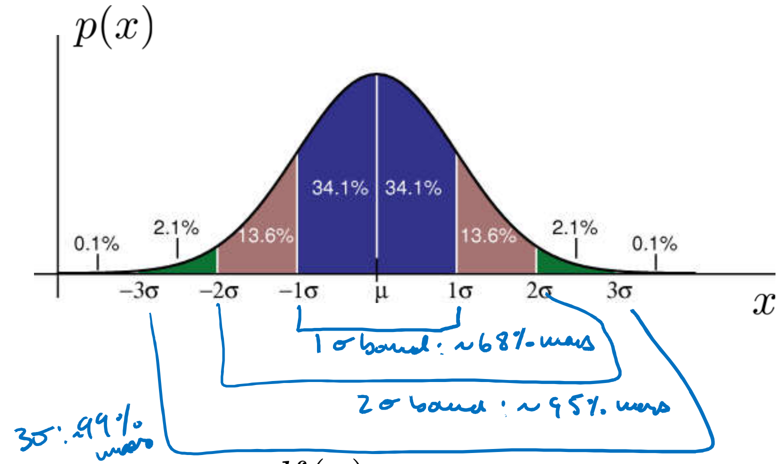


$$\mathcal{N}_{z_\infty}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(z_\infty - \mu)^2\right]$$

- Moreover, if the x_i are **independent and identically distributed (i.i.d.)** with zero mean and some finite variance σ^2, then $p(z_n) \to \mathcal{N}(0, \sigma^2) \quad \text{as } n \to \infty$

The Gaussian (or Normal) Distribution

- The continuous scalar random variable X with realizations x is normally distributed (has a Gaussian distribution) if its pdf is



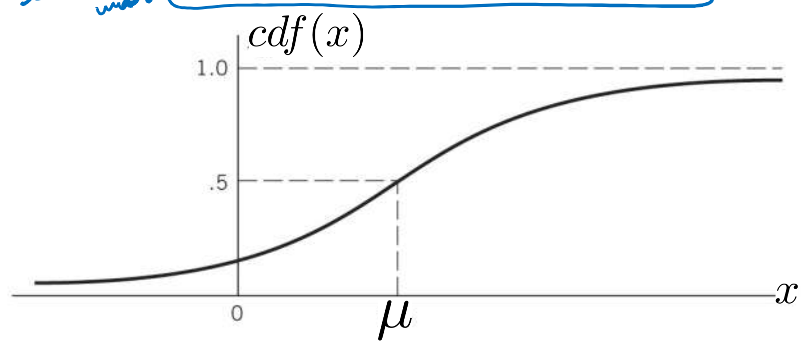
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] = \mathcal{N}(\mu, \sigma^2)$$

$x \sim \mathcal{N}(\mu, \sigma^2)$

- This pdf is <u>completely</u> defined by:

$$\text{mean} = E[x] = \int_{-\infty}^{\infty} x p(x) dx = \mu$$

$$\text{var(x)} = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x - \mu^2) p(x) dx = \sigma^2$$



- Often we will write $x \sim \mathcal{N}(\mu, \sigma^2)$

Gaussians Are Your Friends!

- As we will see throughout course, Gaussians have many exceedingly useful properties
  - Completely described by mean and variance (1st two moments)
  - Single maximum peak at the mean (unimodal), symmetric
  - Show up naturally in physical systems (consequence of Central Limit Theorem)
  - Expectation operations generally easy (especially with linear functions/dynamics)
  - Marginals of multivariate Gaussians are just univariate Gaussians
- Some useful Matlab functions
  - `normpdf`: to compute the pdf of $N(\mu,\sigma)$ at some value x
  - `normcdf`: to compute the cdf of $N(\mu,\sigma)$ at some value x
  - `randn`: to draw samples of standard normal random variables $x \sim N(\underline{0,1})$

to get random samples $s_i \sim N(\mu, \sigma)$?

- Step 1: draw $x_i \sim N(0,1)$ [e.g. x_i = randn]
- Step 2: compute $s_i = \mu + (\sigma \cdot x_i)$

→ can easily check: $E[s_i] = E[\mu + \sigma x_i] = E[\mu] + E[\sigma \cdot x_i] = \underbrace{E[\mu]}_{=\mu} + \sigma E[\overset{0}{\cancel{x_i}}] = \mu$ ✓

$\mathrm{var}(s_i) = E[(s_i - \mu)^2] = (\ldots \text{do yourself} \ldots) \overset{\checkmark}{=} \sigma^2$

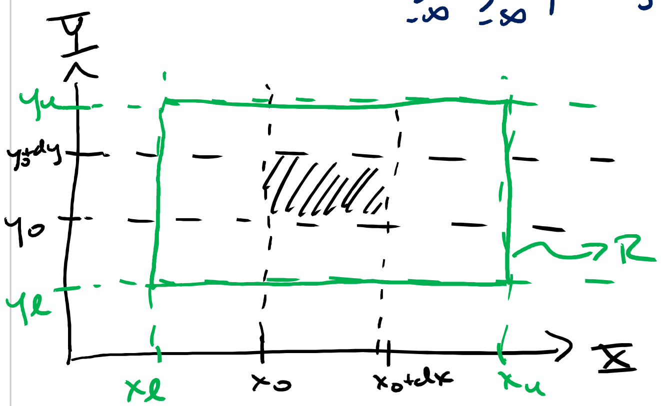
Joint Continuous Random Variables

- How to define joint probabilities for 2 or more continuous RVs?
- Need ***joint probability density functions*** (**joint pdfs**; also called **multivariate pdfs**)

Joint pdf for X & Y : $p_{XY}(x,y) = p(x,y) = p(x \,\&\, y)$: takes in 2 real #'s x & y & returns a scalar value

Such that $P(\{x_0 \le X \le x_0+dx\} \,\&\, \{y_0 \le Y \le y_0+dy\}) = p(x_0,y_0)\cdot dx\,dy$

& $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\,dx\,dy = 1$



↳ Probability that (x,y) lies inside region R: $[x_\ell,x_u]\times[y_\ell,y_u]$

$$= P(x \,\&\, y \in R) = \iint_R p(x,y)\,dx\,dy$$

$$= \int_{x_\ell}^{x_u}\int_{y_\ell}^{y_u} p(x,y)\,dx\,dy$$

→ can also define joint cdf :

$$cdf(x_0,y_0) = C(x_0,y_0) = P(\{x \le x_0\} \,\&\, \{y \le y_0\}) = \int_{-\infty}^{x_0}\int_{-\infty}^{y_0} p(x,y)\,dx\,dy$$

Marginal pdfs and Conditional pdfs

- As with joint probability tables discussed earlier, joint pdfs tell “the whole story”
- Analogous expressions exist for marginalization, conditioning, Bayes’ rule, independence
- **Basically just need to replace summations with integrals**

- Marginal pdf: $p(x) = \int_{-\infty}^{\infty} p(x,y)\,dy$ [marginal pdf for x: "averaging" p(x,y) w.r.t. y]
  $p(y) = \int_{-\infty}^{\infty} p(x,y)\,dx$ [marginal pdf for y: "averaging" p(x,y) w.r.t. x]

- Conditional pdf: $p(x \mid y = \underset{\substack{\uparrow \\ \text{some fixed value for } y}}{\bar{y}}) \triangleq \dfrac{p(x, y=\bar{y})}{p(y=\bar{y})} = \dfrac{p(x, y=\bar{y})}{\int_{-\infty}^{\infty} p(x, y=\bar{y})\,dx}$ [likewise: $p(y \mid x=\bar{x}) = \dfrac{p(x=\bar{x}, y)}{\int_{-\infty}^{\infty} p(x=\bar{x}, y)\,dy}$]

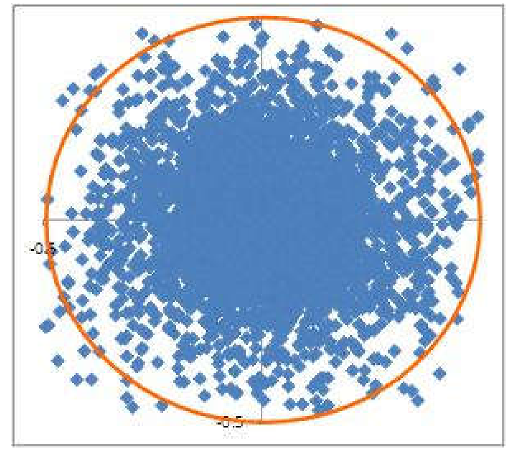
- Bayes' Rule for pdfs (invert cond. pdfs): $p(x \mid y=\bar{y}) = \dfrac{p(x) \cdot p(y=\bar{y} \mid x)}{p(y=\bar{y})} = \dfrac{p(x) \cdot p(y=\bar{y} \mid x)}{\int_{-\infty}^{\infty} p(x)\, p(y=\bar{y} \mid x)\,dx}$

- Independence: Cont. RVs x & y are $\perp\!\!\!\perp$ if & only if $p(x,y) = p(x) \cdot p(y)$ $\forall$ x & y values

(*) All of the above extend to n-dimensional pdfs $p(x_1, x_2, \ldots, x_n)$ in obvious ways

Example: Simple Bivariate Gaussian

- Suppose people throw darts at an infinite board with an x-y coordinate system
- Coordinate (x,y) of each dart hole is continuous 2-dim RV



Let's assume x & y are $\perp\!\!\!\perp$ Gaussian RV's with

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left\{ -\frac{(x^2)}{2\sigma^2} \right\}$$

$$p(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left\{ -\frac{(y^2)}{2\sigma^2} \right\}$$

$\Big\}$ $\sigma_x = \sigma_y = \sigma$

$\rightarrow p(x,y) = p(x) \cdot p(y)$ [since x & y are assumed $\perp\!\!\!\perp$] (assumed by def.)

$$\rightarrow p(x,y) = \left(\frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} \right)\left(\frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{y^2}{2\sigma^2}} \right)$$

$$= \boxed{\frac{1}{2\pi\sigma^2} \cdot \exp\left\{ -\frac{(x^2+y^2)}{2\sigma^2} \right\} = p(x,y)}$$

returns a single scalar value for any pair x&y