02/02/04 Probability theory for random variables with uncountable (i.e. continuous) state space.

Reference: Kloeden + Platen Sec. 1.2, 2.1
Karlin + Taylor Sec. 1.1

Start with state space $S \subseteq \mathbb{R}$ for a random variable X.

How describe X?

Generally: $P_X(B) = \text{Prob}(X \in B)$

where $B \in \mathcal{B}$ (σ-algebra of measurable (Borel) sets)

or

cumulative distribution function

$$F_X(x) = \text{Prob}(X \le x) \quad \text{for } x \in \mathbb{R}$$

If P_X is "absolutely continuous" with respect to Lebesgue measure, then can equivalently describe X by a 'probability density' $p(x)$ such that

$$\text{Prob}(X \in B) = \int_B p(x)\,dx \quad \text{for } B \in \mathcal{B}$$

$$F_X(x) = \int_{-\infty}^{x} p(x')\,dx'$$

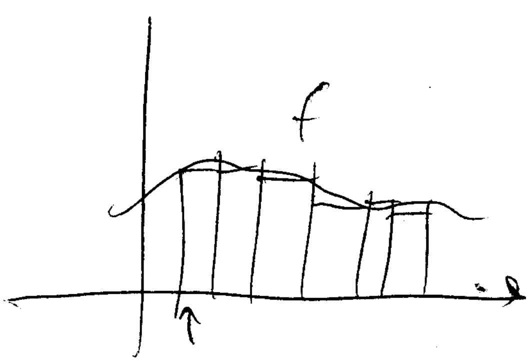
Kloeden + Platen p. 6 typo in first line Eq. (2.4)

- they actually use same defn as in lecture.

How compute averages involving random variables $\mathbb{X}$ with state space $S \subseteq \mathbb{R}$.

If $\mathbb{X}$ has a prob. density $p(x)$ ~~then~~ ~~and~~ and f is a ~~smooth~~ continuous function

then $$\langle f(\mathbb{X}) \rangle = \int_S f(x)\, p(x)\, dx$$



$$\approx \sum f(x_i) \underbrace{p(x_i)\, \Delta x}_{}$$

like discrete probability

$\text{Prob}\left(|\mathbb{X} - x_i| \le \frac{\Delta x}{2}\right)$

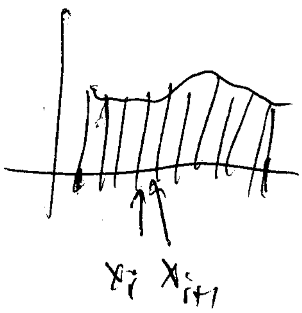
For example $\langle \mathbb{X} \rangle = \int_S x\, p(x)\, dx$

$$\langle \mathbb{X}^2 \rangle = \int_S x^2 p(x)\, dx, \text{ etc.}$$

How generalize if the probability density does not exist?

More generally, for continuous f:

$$\langle f(X) \rangle = \int_S f(x)\, dF_X(x) \quad \text{(Stieltjes integral)}$$



$$= \lim_{\|P\| \to 0} \sum_i f(x_i)\left(F_X(x_{i+1}) - F_X(x_i)\right)$$

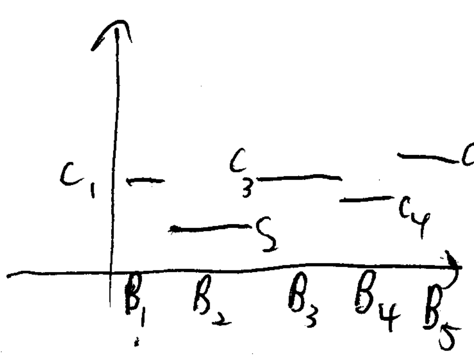
Or can use Lebesgue integral (Kloeden + Platen Sec. 2.2)

$$\langle f(X) \rangle = \int_S f(x)\, dP_X(x)$$ interpreted as a Lebesgue integral.

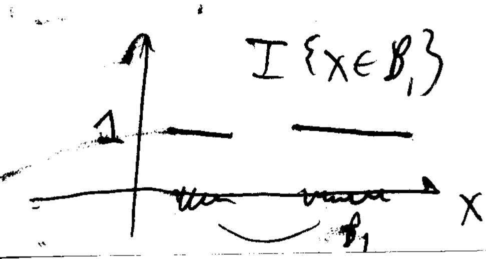
<u>Roughly speaking</u>: Consider "simple functions"

$$s(x) = \sum_{j=1}^{n} c_j\, I\{x \in B_j\}$$

with $c_j \in \mathbb{R}$ and $B_j \in \mathcal{B}$

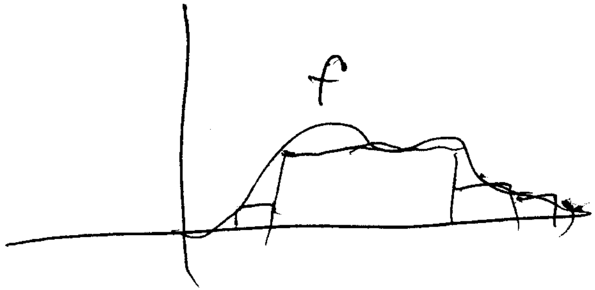


$$\int_S s(x)\, dP_X(x) = \sum_{j=1}^{n} c_j\, P_X(B_j)$$



For nonnegative functions f,

$$\int_S f(x)\, dP_X(x) = \sup_{s \le f} \int_S s(x)\, dP_X(x)$$



For general measurable f, write $f = f_+ + f_-$

HW Prob. 1.2: Prove $\int_S f(x)\, dP_X(x) = \sum_{j=1}^{n} a_j P_j$

May have another problem about Stieltjes integrals

Multiple random variables (or random vectors)

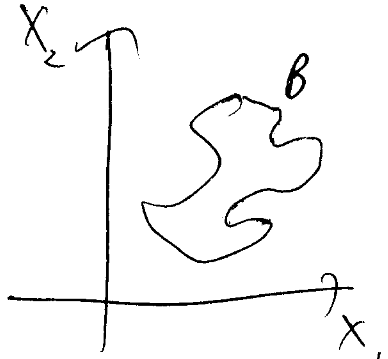
Think of random variables X, Y, Z

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \qquad X \in \mathbb{R}^n$$

Joint probability distributions

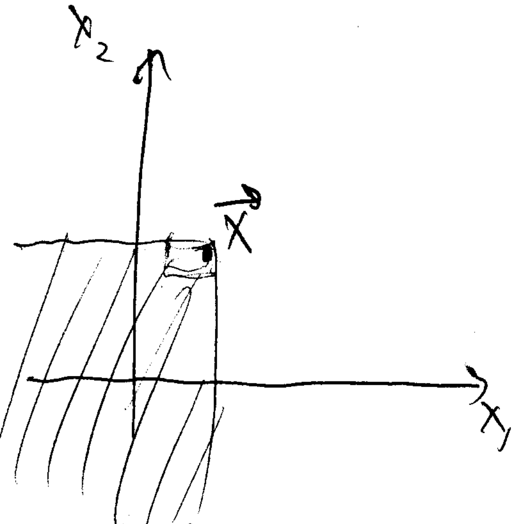
$$\text{Prob}(\vec{X} \in B) \equiv P_X(B) \quad \text{for } B \in \mathcal{B}$$

where $\mathcal{B}$ is σ-algebra of (Borel) measurable subsets of $\mathbb{R}^n$



Cumulative distribution fn exists but more awkward to work with

$$F_X(\vec{x}) = \text{Prob}(X_1 \leq x_1, X_2 \leq x_2, \ldots \text{ and } X_n \leq x_n)$$



If probability density $p(\vec{x})$ exists, then

$$\text{Prob}(X \in B) = \int_B p(\vec{x})\, d\vec{x}$$

for $B \in \mathcal{B}$

$F_X(\vec{x})$ measure P_X of this set.

Computing averages:

$$\langle f(\vec{X})\rangle = \int_S f(\vec{x})\, p(\vec{x})\, d\vec{x}$$ if prob. density exists.

$$= \int_S f(\vec{x})\, dF_{\mathbf{X}}(\vec{x}) \quad (?!)$$ as a Stieltjes integral

$$= \int_S f(\vec{x})\, dP_{\mathbf{X}}(\vec{x})$$ as a Lebesgue integral.

Relationships between continuous random vars.

Consider random variables $\mathbf{X}$ and $\mathbf{Y}$

$\mathbf{X}$ $\downarrow$ pdf $p_{\mathbf{X}}(x)$

$\mathbf{Y}$ $\searrow$ pdf $p_{\mathbf{Y}}(y)$

$$\text{Prob}(\mathbf{X} \in A) = \int_A p_{\mathbf{X}}(x)\, dx \qquad A \in \mathcal{B}_{\mathbf{X}}$$

$$\text{Prob}(\mathbf{Y} \in B) = \int_B p_{\mathbf{Y}}(y)\, dy \qquad B \in \mathcal{B}_{\mathbf{Y}}$$

Definition of independence:

For any $A \in \mathcal{B}_{\mathbf{X}}$, $B \in \mathcal{B}_{\mathbf{Y}}$,

$$\text{Prob}(\mathbf{X} \in A \text{ and } \mathbf{Y} \in B) = \text{Prob}(\mathbf{X} \in A)\,\text{Prob}(\mathbf{Y} \in B)$$

(even if they don't have prob. densities).

If they do have prob densities:

$$\int_A dx \int_B dy\, \underbrace{p_{\mathbf{X},\mathbf{Y}}(x,y)}_{\text{joint pdf}}\, \cancel{dx\,dy} = \left(\int_A dx\, p_{\mathbf{X}}(x)\right)\left(\int_B dy\, p_{\mathbf{Y}}(y)\right)$$

for any $A \in \mathcal{B}_{\mathbf{X}}$, $B \in \mathcal{B}_{\mathbf{Y}}$

$$\Longleftrightarrow \quad p_{\mathbf{X},\mathbf{Y}}(x,y) = p_{\mathbf{X}}(x)\, p_{\mathbf{Y}}(y)$$

Also, if $\mathbb{X}$ and $\mathbb{Y}$ are independent then $\langle f(\mathbb{X})\, g(\mathbb{Y}) \rangle = \langle f(\mathbb{X}) \rangle \langle g(\mathbb{Y}) \rangle$

How describe relationship between dependent r.v.'s $\mathbb{X}$ and $\mathbb{Y}$?

Covariance: $\langle (\mathbb{X} - \mu_{\mathbb{X}})(\mathbb{Y} - \mu_{\mathbb{Y}}) \rangle$

$$= \int_{S_{\mathbb{X}}} dx \int_{S_{\mathbb{Y}}} dy\ (x-\mu_{\mathbb{X}})(y-\mu_{\mathbb{Y}})\, p_{\mathbb{X},\mathbb{Y}}(x,y)$$

$$\mu_{\mathbb{X}} = \langle \mathbb{X} \rangle = \int_{S_{\mathbb{X}}} x\, p_{\mathbb{X}}(x)\, dx$$

$$\mu_{\mathbb{Y}} = \langle \mathbb{Y} \rangle = \int_{S_{\mathbb{Y}}} y\, p_{\mathbb{Y}}(y)\, dy$$

Correlation coef: $$\rho_{\mathbb{X},\mathbb{Y}} = \frac{\langle (\mathbb{X} - \mu_{\mathbb{X}})(\mathbb{Y} - \mu_{\mathbb{Y}}) \rangle}{\sigma_{\mathbb{X}}\ \sigma_{\mathbb{Y}}}$$

For more details: Conditional probability and expectation.

Can't define

$$\text{Prob}(A \mid Y = y) = \frac{\text{Prob}(A \text{ and } Y = y)}{\text{Prob}(Y = y)}$$

Denominator is often zero.

Have to define in a "weak" way.

Suppose Y has a probability density $p_Y(y)$

Then <u>define</u> $\text{Prob}(A \mid Y = y)$ so that

$$\int_B \text{Prob}(A \mid Y = y)\, p_Y(y)\, dy = \text{Prob}(A \text{ and } Y \in B)$$

for any $B \in \mathcal{B}_Y$.

If $B = \{ y \in S_Y : |y - y_0| \le \varepsilon \}$ then

$$\text{Prob}(A \mid Y = y_0)\, p_Y(y_0)\, 2\varepsilon \approx \text{Prob}(A \text{ and } |Y - y_0| \le \varepsilon)$$

for <u>small ε.</u>

Intuitive: $\text{Prob}(A \mid Y = y_0) \approx \dfrac{\text{Prob}(A \text{ and } |Y - y_0| \le \varepsilon)}{\text{Prob}(|Y - y_0| \le \varepsilon)}$

Billingsley Sec. 33 + 34

Sinai Lecture 4

Conditional expectation?

If the conditional probability has a density:

Suppose $\mathrm{Prob}(\mathbb{X} \in A \mid \mathbb{Y} = y) = \int_A p_{\mathbb{X}|\mathbb{Y}}(x|y)\,dx$ for all $A \in \mathcal{B}_{\mathbb{X}}$

Then $\langle \mathbb{X} \mid \mathbb{Y} = y \rangle \equiv \mathbb{E}[\mathbb{X} \mid \mathbb{Y} = y]$

$$= \int_{S_{\mathbb{X}}} x \, \underline{p_{\mathbb{X}|\mathbb{Y}}}(x|y)\,dx$$

Marginal probability density:

Suppose $\mathbb{X}$ and $\mathbb{Y}$ have a joint probability density $p_{\mathbb{X},\mathbb{Y}}(x,y)$.

Then the marginal probability density for $\mathbb{X}$

$$p_{\mathbb{X}}(x) = \int_{S_{\mathbb{Y}}} dy \; p_{\mathbb{X},\mathbb{Y}}(x,y)$$

All this extends directly to more r.v.'s:

$$p_{\mathbb{X}}(x_1) = \int_{S_{\mathbb{X}}^{n-1}} dx_2 \cdots dx_n \; p_{\vec{\mathbb{X}}}(\vec{x}) \quad \text{for} \quad \mathbb{X} = \begin{pmatrix} \mathbb{X}_1 \\ \vdots \\ \mathbb{X}_n \end{pmatrix}$$

How simulate r.v.'s in continuous state space

Kloeden + Platen Sec. 1.3

Numerical Recipes Ch. 7

Inverse transform method (always works but may be cumbersome)

Consider $S \subseteq \mathbb{R}$

Recall we showed that if U is a uniformly distributed random number on $[0,1]$. Then $F_X^{-1}(U)$ is a random variable w/ same probability distribution as X.

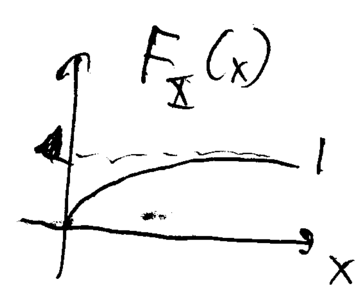
Example: Exponential distribution

PDF: $p(x) = \lambda e^{-\lambda x}$ for $x \geq 0$

$\qquad = 0$ for $x < 0$

CDF: $F_X(x) = 1 - e^{-\lambda x}$ for $x \geq 0$

$\qquad = 0$ for $x < 0$



$$X_{sim} = F_X^{-1}(U) = -\frac{\ln(1-U)}{\lambda}$$

where U is uniform r.v. on $[0,1]$

Gaussian: $p(x) = \dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$

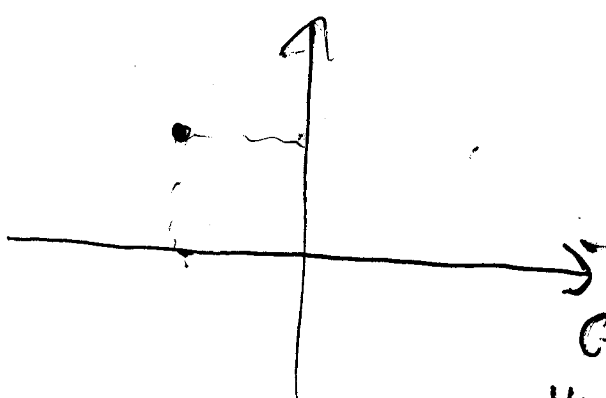
F_X involves erf

Inverse transform method is cumbersome... for this case.

Better: Box-Muller method:

Use the fact that 2 independent Gaussian random variables have, when plotted on a 2-dimensional plane, a uniform distribution in angle and a radial distribution $\propto re^{-r^2}$

use inverse xform method



Generate 2 independent uniform random $U_1, U_2 \in [0,1]$

$$N_1 = \sqrt{-2\ln U_1}\cos 2\pi U_2$$

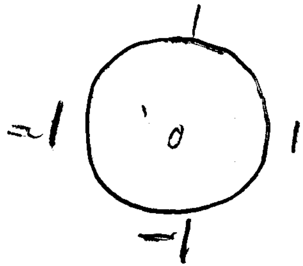
$$N_2 = \sqrt{-2\ln U_1}\sin 2\pi U_2$$

are two independent Gaussian r.v.'s with mean $=0$, variance $=1$.

Even better (faster): Polar-Marsaglia Method

- avoids trig evals

Chooses a random point in unit circle (uniformly).



Let R be the radius and θ be the angle.... to be continued

Homework 1 due 02/12 5 PM