ASEN 5044, Fall 2018

Statistical Estimation for Dynamical Systems

Lecture 10:
Expectation Operator and Expected Values

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

Thurs 9/27/2018

University of Colorado Boulder

Aerospace Engineering

Announcements

- **HW 4 Posted -- Due Thurs 10/4 at 11 am**
- Submit to Canvas

- **Special topic extra lecture: tomorrow (Fri 9/28) at 1-1:50 pm, ECCS 1B12**
  - **Notes will be posted, lecture will be recorded**
  - **All encouraged to attend if possible (PhD students especially)**

- **Midterm 1: will be posted next Thursday 10/4**
  - One week long take home exam posted to Canvas
  - Due Thurs 10/11/2018 on Canvas
  - Open book/notes – honor code applies (must complete by yourself)

Last Time…

- Marginal and conditional probabilities
- Bayes' rule, independence

- discrete and continuous random variables (i.e. "random quantities")

- probability mass functions (pmfs) for discrete random variables

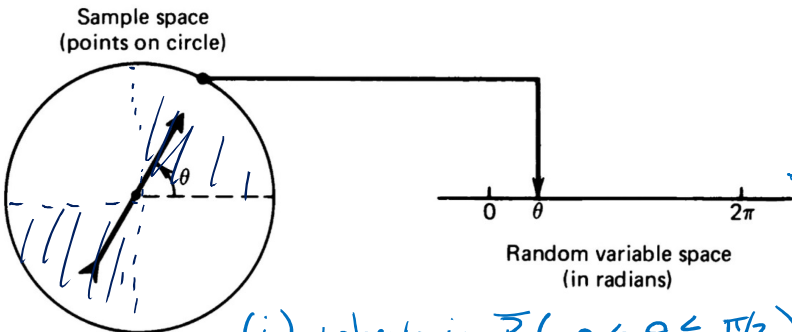
- probability density functions (pdfs) for continuous random variables

↳ events = lengths / ~~&~~ intervals

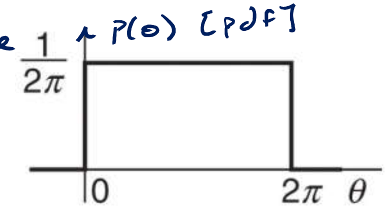
PDF Example: Spinning Pointer on Wheel

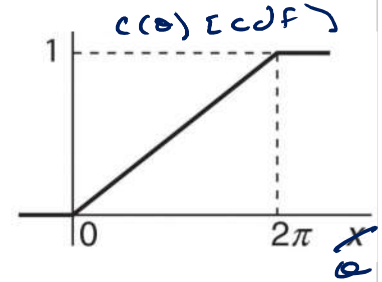
- If spinner fairly constructed, then θ has uniform pdf : $\theta \sim \mathcal{U}[a,b]$, $a = 0$, $b = 2\pi$

$$\mathcal{U}[0,2\pi] = \begin{cases} \frac{1}{2\pi}, & x \in [0,2\pi], \text{ distributed as} \\ 0, & \text{o'wise} \end{cases}$$



$$\rightarrow p(\theta) = \mathcal{U}_\theta[0,2\pi]$$





(i) what is $P(0 \le \theta \le \pi/2) = ?$

$$\rightarrow P(0 \le \theta \le \pi/2) = \int_0^{\pi/2} p(\theta)\,d\theta = \int_0^{\pi/2} \frac{1}{2\pi}\,d\theta = \frac{1}{2\pi}\int_0^{\pi/2} 1\,d\theta = \frac{1}{2\pi}\cdot\theta\Big|_0^{\pi/2} = \boxed{\frac{1}{4}}$$

$\rightarrow$ what is $P([0 \le \theta \le \pi/2] \text{ OR } [\pi \le \theta \le 3\pi/2]) = ?$

b/c intervals (events) are disjoint:

$$= P(0 \le \theta \le \pi/2) + P(\pi \le \theta \le 3\pi/2)$$

(where $P(\pi \le \theta \le 3\pi/2) = \int_\pi^{3\pi/2} \frac{1}{2\pi}\,d\theta = \frac{1}{2\pi}\cdot\theta\Big|_\pi^{3\pi/2} = \frac{1}{4}$)

$$= \frac{1}{4} + \frac{1}{4} = \boxed{\frac{1}{2}}$$

note: $\int_0^{2\pi} p(\theta)\,d\theta = \int_{-\infty}^{\infty} p(\theta)\,d\theta = 1$
[pdf always integrates to 1 over $[-\infty,\infty]$]

Today…

Continue Continuous RVs and Probability Densities

- More on expectation operators and expected values, examples

READ SIMON BOOK, CHAPTER 2.5

Expected Values and the Expectation Operator

- Not surprisingly, the function Y = g(X) of a RV X is also a RV (i.e. Y is a RV)
- We could try to find the distribution of Y, but this can be difficult or unnecessary
- Sometimes we just need a "summary of what to expect" from Y without enumerating all possible values for Y
- i.e. what is the "average value" of some arbitrary function g(x) of random var X?

Discrete Case

$$E[g(x)] = \sum_{i=1}^{N_x} g(x=i)P(x=i)$$

= single # (constant w.r.t x)

Continuous Case

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

= single # (constant w.r.t. x)

$E[\cdot]$ = expectation operator:
returns expected value of g(x) [the argument (·)]

Interpretation of Expected Values

- Consider the "relative frequency" view of probabilities (for discrete RVs):

$$p_i = \lim_{N \to \infty} \frac{N_i}{N}, \quad x_i \in \{1, 2, 3, \ldots, N_x\} \implies \text{\# of times we typically expect to see } x_i \text{ in } N \text{ trials}$$

(prob. of RV x_i for event $i = 1, \ldots, N$ given N_i occurrences in N trials)

as $N \to \infty$:
$N_1 = p_1 \cdot N$ (# typical times $x_i = 1$)
$N_2 = p_2 \cdot N$ (# typical times $x_i = 2$)
$\vdots$
$N_{N_x} = p_{N_x} \cdot N$ (" " $x_i = N_x$)

- So given some sample of outcomes, the 'typical' N-sample mean for corresponding RVs would be:

$$\bar{x}_{sample} = \frac{N_1 \cdot x_1 + N_2 \cdot x_2 + \cdots + N_{N_x} \cdot x_{N_x}}{N} \underset{\text{as } N \to \infty}{\implies} \frac{[p_1 \cdot \cancel{N}] \cdot x_1 + [p_2 \cdot \cancel{N}] \cdot x_2 + \cdots + [p_{N_x} \cdot \cancel{N}] \cdot x_{N_x}}{\cancel{N}}$$

(finite N)

[typical sample mean or sample avg. value]

$$= p_1 \cdot x_1 + p_2 \cdot x_2 + \cdots + p_{N_x} \cdot x_{N_x} = \sum_{i=1}^{N_x} x_i \cdot p_i = E[x_i]$$

No dependence on N!

- The **expected value (EV)** = conceptual average obtained over infinite # of trials N
  - Key idea: don't actually need to run infinite # of trials N if we know probability of outcomes
  - EV is what you expect to see in a "typical" random trial (not what you actually will see – b/c trial is random!)
  - Expected value is NOT the same as the sample average (sample mean) for finite N
  - Expected value says NOTHING about the actual number you will obtain for finite N

Some Common/Important Expected Values

- Things you will see a lot in estimation problems:

(1st moment) Mean or average of RV x: $E[x] = \bar{x} = \mu_x = \begin{cases} \int_{-\infty}^{\infty} x\, p(x)\, dx & \text{(cont. RV)} \\ \sum_{i=1}^{N_x} x \cdot P(x) & \text{(disc. RV)} \end{cases}$

- 2^{nd} moment: $E[x^2] = \begin{cases} \int_{-\infty}^{\infty} x^2 p(x)\, dx & \text{(cont.)} \\ \sum_{i=1}^{N_x} x^2\, P(x) & \text{(disc.)} \end{cases}$

- Variance (2^{nd} moment about the mean): $\mathrm{var}(x) = \sigma_x^2 = \sigma^2 = E[(x-\mu)^2] > 0$

$$= \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^2 \cdot p(x)\, dx \\ \sum_{i=1}^{N_x} (x-\mu)^2 \cdot P(x) \end{cases}$$

$\updownarrow$

Standard deviation $= \mathrm{std}(x) = \sigma_x = \sigma = \sqrt{\mathrm{var}(x)} > 0$

- Higher order moments: $E[x^n] = \begin{cases} \int_{-\infty}^{\infty} x^n p(x)\, dx \\ \sum_{i=1}^{N_x} x^n\, P(x) \end{cases}$ [n^{th} moment] } tells us shape info of dist. fcn (e.g. skewness, kurtosis, ...)

- Expected reward/cost function $J(x)$: $E[J(x)] = \begin{cases} \int_{-\infty}^{\infty} J(x)\, p(x)\, dx \\ \sum_{i=1}^{N_x} J(x)\, P(x) \end{cases}$

Example #1: Die Rolls

(a) Compute the expected face value i for the roll of a single fair die (mean roll value)

$$\bar{x} = \mu_x = E[x] = \sum_{i=1}^{6} x_i \cdot P(x_i) = 1 \cdot P(x_i = 1) + 2 \cdot P(x_i = 2) + \dots + 6 \cdot P(x_i = 6)$$

$$= 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + \dots + 6 \cdot \tfrac{1}{6} = \tfrac{1}{6} \cdot 21$$

$$= \boxed{\bar{x} (= \mu = E[x]) = 3.5}$$

(b) Find expected reward ("expected take") for a single roll, if given reward function R(i)

i	$R(i)$ (\$)
1	0
2	5
3	10
4	10
5	5
6	0

$$\rightarrow E[R] = \sum_{i=1}^{6} R(i) P(i)$$

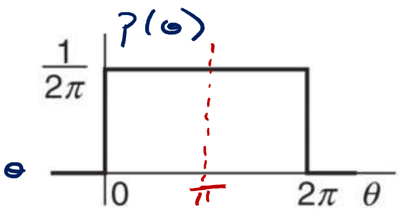
$$= R(i=1) \cdot P(i=1) + R(i=2) \cdot P(i=2) + \dots + R(i=6) \cdot P(i=6)$$

$$= 0 \cdot \tfrac{1}{6} + 5 \cdot \tfrac{1}{6} + 10 \cdot \tfrac{1}{6} + 10 \cdot \tfrac{1}{6} + 5 \cdot \tfrac{1}{6} + 0 \cdot \tfrac{1}{6}$$

$$= (2) \cdot 5 \cdot \tfrac{1}{6} + (2) \cdot 10 \cdot \tfrac{1}{6} = \tfrac{30}{6} = \boxed{\$5} = E[R]$$

Example #2: Moments of Uniform PDF

- Find mean, 2[nd] moment and variance of θ for spinning wheel problem



$$p(\theta) = \mathcal{U}_\theta[0, 2\pi] = \begin{cases} \frac{1}{2\pi}, & \theta \in [0, 2\pi] \\ 0, & \text{o'wise} \end{cases}$$

$\rightarrow$ mean $= \bar{\theta} = \mu_\theta = E[\theta] = \int_{-\infty}^{\infty} \theta \cdot p(\theta)\, d\theta = \int_{-\infty}^{\infty} \theta \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta$

$\rightarrow$ use definition of $\mathcal{U}_\theta[0, 2\pi]$: $E[\theta] = \int_0^{2\pi} \theta \cdot \frac{1}{2\pi}\, d\theta$ (b/c $p(\theta)$ is zero for $\theta \notin [0, 2\pi]$)

$$= \frac{1}{2\pi} \cdot \left[\frac{\theta^2}{2} \right|_0^{2\pi} \Rightarrow \boxed{\bar{\theta} = \pi}$$

$\rightarrow$ 2[nd] moment : $E[\theta^2] = \int_{-\infty}^{\infty} \theta^2 \cdot p(\theta)\, d\theta = \int_{-\infty}^{\infty} \theta^2 \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta$

use definition of $\mathcal{U}_\theta[0, 2\pi]$ & fact that $p(\theta) = 0$ for $\theta \notin [0, 2\pi]$ $\longrightarrow$ $E[\theta^2] = \int_0^{2\pi} \theta^2 \cdot \frac{1}{2\pi}\, d\theta = \frac{1}{2\pi} \int_0^{2\pi} \theta^2\, d\theta = \frac{1}{2\pi} \cdot \left[\frac{\theta^3}{3} \right|_0^{2\pi} = \frac{(2\pi)^2}{3}$

$$\rightarrow \boxed{E[\theta^2] = \frac{(2\pi)^2}{3}}$$

Example #2: Moments of Uniform PDF (cont'd)

$$\text{Var}(\theta) = E[(\theta - \bar{\theta})^2] = \int_{-\infty}^{\infty} (\theta - \bar{\theta})^2 \cdot p(\theta)\, d\theta \qquad \left(\bar{\theta} = E[\theta] = \text{const. w.r.t. } \theta = \pi\right)$$

$$= \int_{-\infty}^{\infty} (\theta - \bar{\theta})^2 \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta$$

$$= \int_{-\infty}^{\infty} (\theta^2 - 2\theta \cdot \bar{\theta} + \bar{\theta}^2) \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta$$

$$= \underbrace{\int_{-\infty}^{\infty} \theta^2 \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta}_{= E[\theta^2]} - 2\bar{\theta} \underbrace{\int_{-\infty}^{\infty} \theta \cdot \mathcal{U}_\theta[0, 2\pi]\, d\theta}_{= E[\theta] = \bar{\theta}} + \bar{\theta}^2 \underbrace{\int_{-\infty}^{\infty} \mathcal{U}_\theta[0, 2\pi]\, d\theta}_{= 1}$$

$$= E[\theta^2] - 2\bar{\theta} \cdot \bar{\theta} + \bar{\theta}^2 = E[\theta^2] - 2\bar{\theta}^2 + \bar{\theta}^2$$

$$= E[\theta^2] - \bar{\theta}^2$$

$$= \boxed{E[\theta^2] - (E[\theta])^2 = \text{var}(\theta)}$$

$$\boxed{= \frac{\pi^2}{3}} \quad (\text{using } E[\theta^2] \text{ \& } \bar{\theta} \text{ values from before}$$

Useful Properties of Expectations (both continuous and discrete)

- FACT 1: The expectation operator is linear

$$E_x[\alpha f(x) + \beta g(x)] = \alpha \cdot E_x[f(x)] + \beta \cdot E_x[g(x)]$$

for any constants α & β & integrable $f(x)$ & $g(x)$

$$\& \; E_x[(\cdot)] = \begin{cases} \int_{-\infty}^{\infty} (\cdot)\, p(x)\, dx & [\text{for cont. RVs}] \\ \sum_{i=1}^{N_x} x_i \, \mathcal{I}(x_i) \ldots & [\text{for disc. RVs}] \end{cases}$$

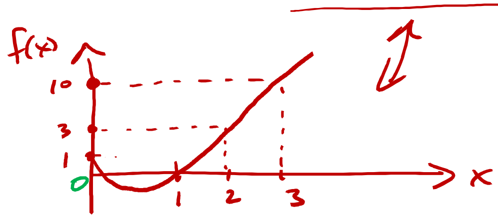
- FACT 2: Variance can always be computed more simply as

$$E[(x-\bar{x})^2] = \text{var}(x) = E[x^2] - (E[x])^2$$

$$\left(= E[x^2] - \mu_x^2 \right)$$

Example #3: Expected Values of Functions

Find E[f(x)] if f(x) = $2x^2$ -3x +1 for the skewed pdf p(x) $= \begin{cases} \frac{2}{9}(3-x), & \text{for } 0 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$



Note: $f(\bar{x}) = f(1) = 0$

can show:

$E[x] = \bar{x} = 1$

$E[x^2] = 1.5$

$$\rightarrow E[f(x)] = E[2x^2 - 3x + 1]$$

(by linearity of E[·]) $= 2 \cdot E[x^2] - 3 \cdot E[x] + E[1]$

$$= 2 \cdot (1.5) - 3 \cdot (1) + E[1] \rightarrow \text{but } E[1] = \int_{-\infty}^{\infty} 1 \cdot p(x)\,dx = 1$$

$$= 2(1.5) - 3 + 1$$

$$= 3 - 3 + 1 = \boxed{1 = E[f(x)]} \neq f(E[x]) = 0$$