Probability Theory - II

EE698V - Machine Learning for Signal Processing

Vipul Arora

Announcements

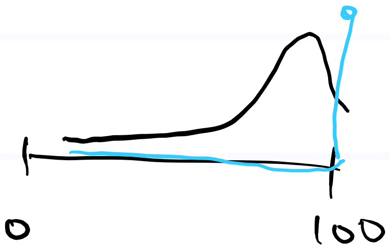
- Quiz on Saturday 8-9PM in L-7
- Should we have a discussion hour tomorrow?
    - 12-1PM in ACES. Others can come 1-2PM in my office.
- If you are facing any problems with any aspect of the course, please do let me know

References

- PRML Section: 1.2 (**highly recommended**)

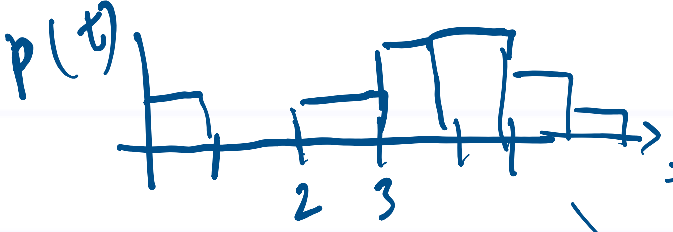
Intuitive Understanding

- Bayesian view
  - My belief on the value of a random variable

    m
  - E.g.,
    - How many marks will I get in the exam?
    - How long will I take to complete my assignment? → t
- You get lumps of samples; approximate with a function

Examples



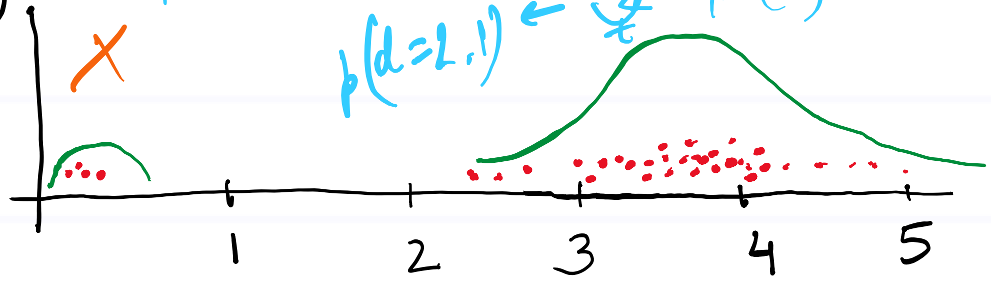
$t \to$ r.v. (cont.)

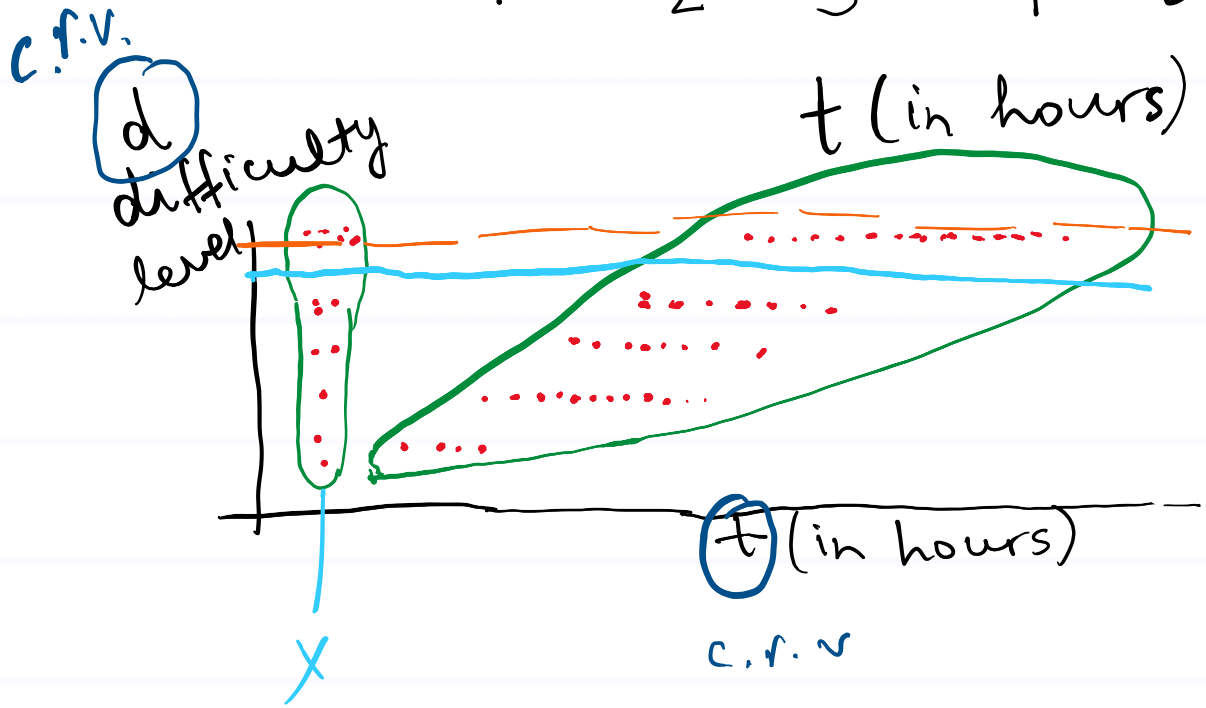
(2.1, 2.5)

$$p(t \mid d=2.1) = \frac{p(t, d=2.1)}{\int dt\, p(t, d=2.1)}$$

$p(d=2.1) \leftarrow$

- How long shall I take to complete this assignment?
- How long do people take to complete assignments ...
- $p(t,d)$
- Limited observation capacity, so we use intuition to build models



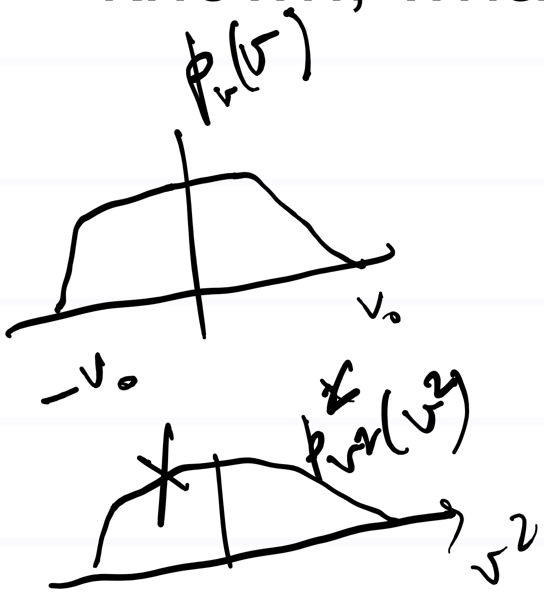


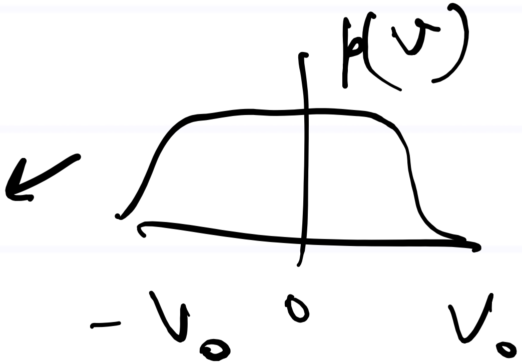
Marginalization, Conditional prob., etc.

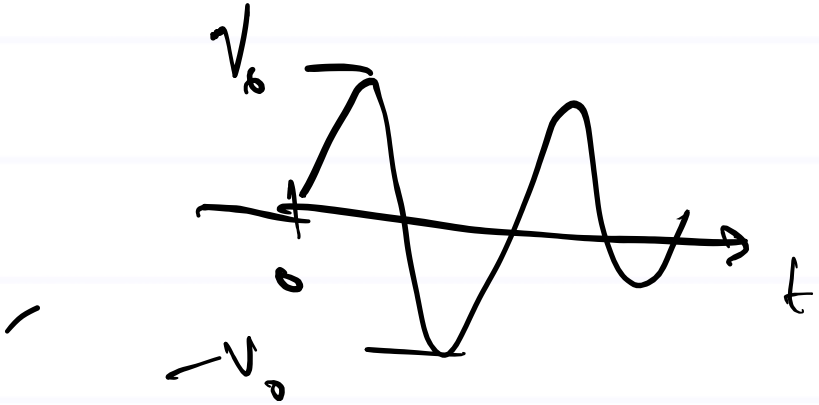
- $p(x) = \int p(x, y) dy$ Sum Rule
- $p(x, y) = p(x|y)p(y)$ Product Rule
- $p(y|x) = \frac{p(x|y)p(y)}{p(x)}$ Bayes Rule

Transformation of Variables

- Given $p(x)$, what is $p(x^2)$?
- For voltage readings across a resistor, p(voltage) is known, what is p(power)?



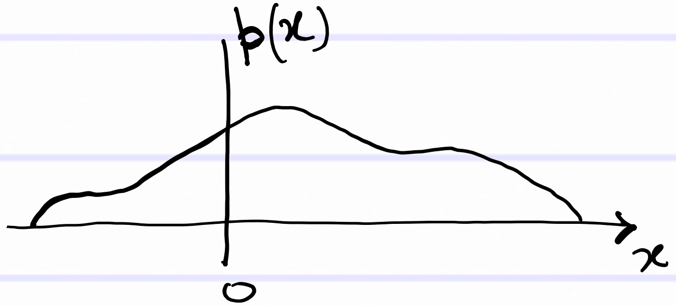


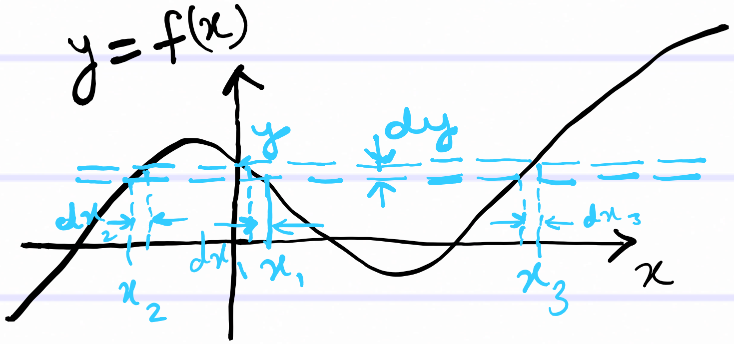


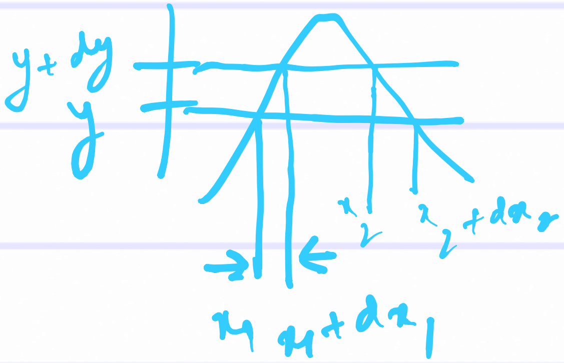
$$\text{power} = \frac{v^2}{R}$$

Transformation of Variables

- Given $p_x(x)$ and $y = f(x)$, find $p_y(y)$
- At some x_1, $y = f(x_1)$
- But there may be other x too that may to same y, say $f(x_i) = y$ for $i = 1,2,\ldots,N$
- $\underbrace{p_y(y)|dy|}_{y \in (y,\, y+dy)} = \underbrace{p_x(x_1)|dx_1|}_{x \in (x_1,\, x_1+dx_1)} + p_x(x_2)|dx_2| + \cdots$
- $p_y(y) = \sum_i \frac{p_x(x)}{\left|\frac{dy}{dx}\right|}\bigg|_{x_i}$







$dx_2 \neq dx_1$

Expectations

- What is the expected (average) width of a rose? How much do they deviate from this average value?

$A = \pi r^2$

- What is the expected area of the top of a rose?

$\underset{A}{p}(\pi r^2)$

- What is the expected (average) width of all roses with height greater than 2cm?

w, h

$E[\cancel{p}(w \mid h > 2)]$

Expectations

- Expected value of f(x) — Mean=μ
  - $\mathbb{E}[f(x)] = \int f(x)p(x)dx$
- Variance of f(x) — Variance=σ^2
  - $\mathrm{var}[f(x)] = \mathbb{E}[\,(f(x) - \mathbb{E}[f(x)])^2\,]$ — Standard Deviation=σ

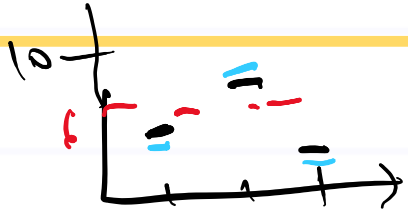
Conditional Expectation

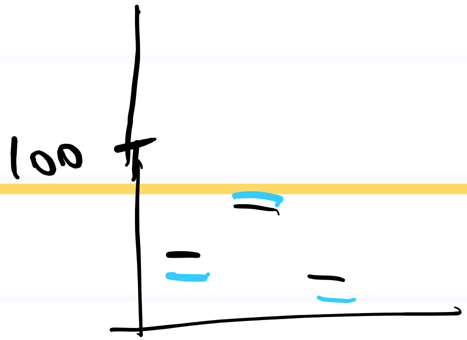
- Conditional Expectation
  - $\mathbb{E}_x[f(x)|y] = \int f(x)p(x|y)dx$
- Covariance
  - $\mathrm{cov}[\boldsymbol{x}, \boldsymbol{y}] = \mathbb{E}[(\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}])(\boldsymbol{y}^T - \mathbb{E}[\boldsymbol{y}^T])]$

a

d

Correlation





- $\rho_{x,y} = \text{corr}(x, y) = \dfrac{\text{cov}[x,y]}{\sigma_x \sigma_y}$

$\text{cov} = 5$

$\text{cov} = E[(x-\mu_x)(y-\mu_y)]$

$\sum_i (x-6)(y-6)$

$= \sum_i (x-60)(y-60)$

$=$

- Uncorrelated $\Rightarrow \quad \rho_{x,y} = 0$
- Prove that two independent random variables are always uncorrelated
- Are two uncorrelated random variables always independent? Prove your answer

$$p(x, y) = p(x)\, p(y)$$

$$\text{corr} = \frac{1}{\sigma_x \sigma_y} E\left[(x-\mu_x)(y-\mu_y)\right]$$

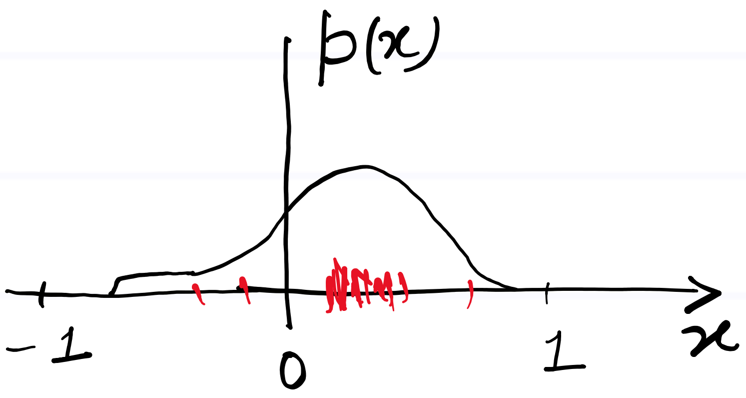
$$\text{num.} = \iint (x-\mu_x)(y-\mu_y)\, \underbrace{p(x, y)}_{p(x)\, p(y)}\, dx\, dy$$

$$= \underbrace{\int (x-\mu_x)\, p(x)\, dx}_{\mu_x - \mu_x} \quad \underbrace{\int (y-\mu_y)\, p(y)\, dy}$$

Samples vs pdf

- Samples $\{s_1, s_2, s_3, \ldots, s_{50}\}$ obtained by measuring x
- $E[x] = \int x\, p(x)dx \approx \sum_i x_i\, p(x_i)$
- $E[x] \approx \frac{1}{50}\sum_i s_i$

No p(x) here



$\int x\, p(x)\, dx$

$\cancel{E[x] = \sum_i s_i\, p(s_i)}$

S are samples from discr. r.v. X

Samples vs pdf

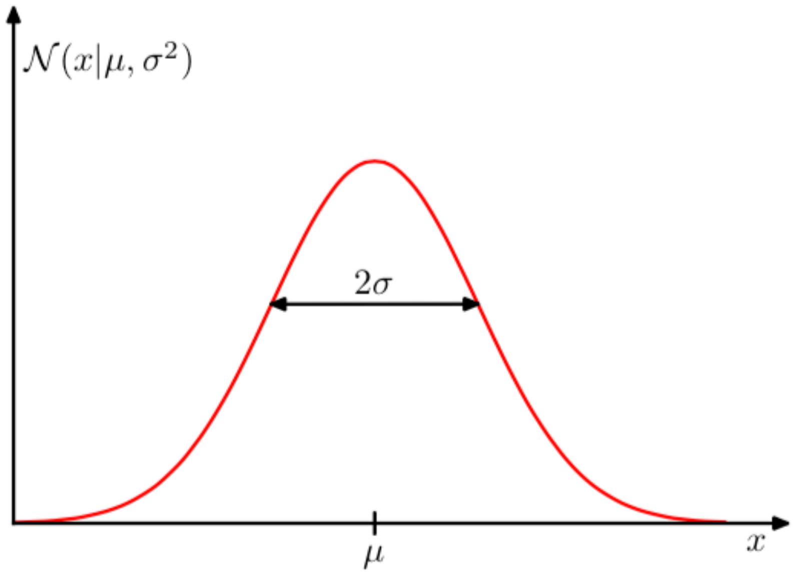
- $E[f(x)] = \int f(x)p(x)dx$
- $E[f(x)] \approx \frac{1}{50}\sum_i f(s_i)$

Standard distributions

- Known or easy-to-compute properties. E.g.,
  - integration
  - marginalization
  - derivatives
- Examples
  - Bernoulli, Multinomial, etc.
  - Gaussian, Laplacian, etc.

Gaussian Distribution

$$\mathcal{N}\left(x|\mu,\sigma^2\right) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

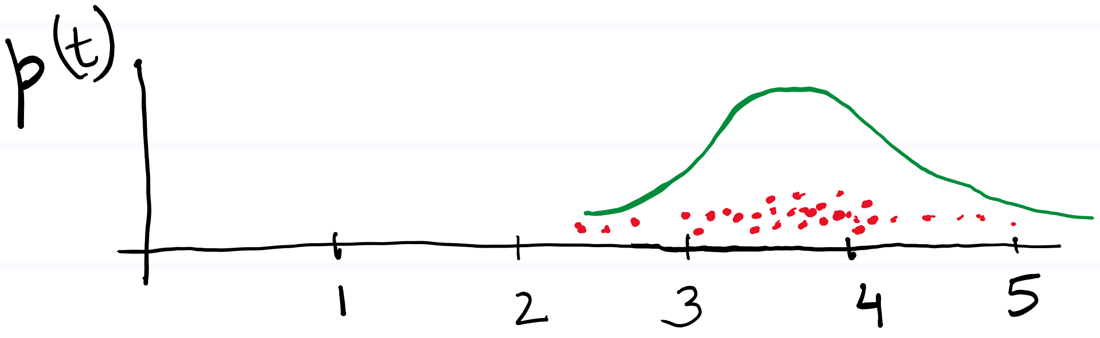


$$\mathbb{E}[x] = \int_{-\infty}^{\infty} \mathcal{N}\left(x|\mu,\sigma^2\right) x \,\mathrm{d}x = \mu.$$

$$\mathbb{E}[x^2] = \int_{-\infty}^{\infty} \mathcal{N}\left(x|\mu,\sigma^2\right) x^2 \,\mathrm{d}x = \mu^2 + \sigma^2.$$

Estimating the parameters

- I have samples. I want to model it with Gaussian distribution. How do I estimate the model parameters?



- Random variable is t
- I want to maximize the probability of observed samples $\{s_1, s_2, \dots, s_N\}$