Supervision Notes 3F3

Krish Shah xxxx@cam.ac.uk

Engineering Department, Cambridge

This document is written for my own reference!

PDFs and PMFs

- be comfortable with the distinction between PMFs and PDFs
- understand eg. a dice roll can be written as

$$f_X(x) = \sum_{i\in\{1,\ldots 6\}} \frac{1}{6}\delta(x-i)$$

Intuition for the change of variables

- Sketch graphs to derive

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{x-a}{a}\right)$$

for $Y = aX + b$

- spot the link to formulas

$$f_X(x) = \frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right)$$

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Multiplication and Convolution

in these cases, explain how the random variable Z relates to X and Y:

- $f_Z(z) = [f_X \star f_Y](z)$
- $f_Z(z) = f_X(z) \times f_Y(z)$
- $F_Z(z) = [F_X \star F_Y](z)$
- $F_Z(z) = F_X(z) \times F_Y(z)$

as an extension, you can also reason through that $Z = X + Y$ corresponds to $\int f_X(z-u) f_Y(u) du$, and derive similar expressions for $Z = XY$

Fun puzzle

Find $\mathbb{E}[\min(X,Y)]$ when $X, Y \sim \text{Uniform}(0,1)$

- a really clean approach that uses $F_Z(z) = F_X(z) \times F_Y(z) \Leftrightarrow Z = \min(X,Y)$
- a messier approach integrating $\iint \min(x,y)$
- a geometric approach where $\min(x,y)$ is a pyramid $\Rightarrow$ the volume is $\frac{1}{3}$ and the probability is $\frac{1}{3}$

Big Picture

$$\{1\ldots 6\} \longrightarrow \mathbb{N} \longrightarrow \mathbb{R} \longrightarrow \overbrace{\mathbb{R}^n \longrightarrow \mathbb{R}^{\mathbb{N}}}^{\text{3F3}} \longrightarrow \underbrace{\mathbb{R}^{\mathbb{R}}}_{\text{4F7}}$$

(note that $\mathbb{R}^{\mathbb{N}}$ is also denoted $\mathbb{R} \times \mathbb{N}$)

- 3F3 looks at simpler instances of $\mathbb{R}^{\mathbb{N}}$ with WSS, ergodicity, strictly stationary ect.
  - WSS is (1) constant mean over X_n (2) finite variance (3) autocorrelation function
  - strictly stationary is that every segment looks the same

Cauchy Distribution

What does it mean for a distribution to have infinite variance?

- 'bad footballer': kick a football forwards at an angle $\varphi \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$ and try work out the expected x value
- pdf looks like $\frac{1}{1+x^2}$ and cdf looks like $\frac{\arctan(x)}{\pi} + \frac{1}{2}$
- expected variance is infinite, despite mean is finite

Wiener Filter

- can you draw the picture, d_n, v_n, h_n, x_n, $\hat{d}_n$, ε_n
- what does it mean for a filter to be optimal
  - h_n that minimizes $\mathbb{E}[\varepsilon_n^2]$
  - x_k uncorrelated with ε_k
- what does it mean for two things to be uncorrelated?

Multivariate Calculus Big Picture

- Differentiating scalar y w.r.t. scalar x is $\frac{\mathrm{d}y}{\mathrm{d}x}$
- Differentiating scalar y w.r.t *vector* $\boldsymbol{x}$ is $\frac{\partial}{\partial \boldsymbol{x}}(y)$ and yields a vector with the shape of $\boldsymbol{x}$. It is analogous to the vector gradient ∇y.

- Differentiating a *vector* $\boldsymbol{y}$ w.r.t a *scalar* x is applied element wise.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\boldsymbol{y}) = \begin{pmatrix} \frac{\mathrm{d}y_0}{\mathrm{d}x} \\ \vdots \end{pmatrix} \text{ for each element in } \boldsymbol{y}$$

Multivariate Cheat Sheet

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{b}) = \boldsymbol{b}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{b}^T \boldsymbol{x}) = \boldsymbol{b}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{B}\boldsymbol{x}) = (\boldsymbol{B} + \boldsymbol{B}^T)\boldsymbol{x}$$

It's quite rewarding to reason through $\frac{\partial}{\partial x_i}(\boldsymbol{B}\boldsymbol{x})$ and use that to build up to deriving $\frac{\partial}{\partial x_i}(\boldsymbol{x}^T \boldsymbol{B}\boldsymbol{x})$ using the product rule.

Normal Distribution Cheat Sheet

$\mathcal{N}(x; \mu, \sigma^2)$ is the value of the Gaussian *evaluated* at x. This notation makes algebra easy, eg.

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

$$\mathcal{N}(x; \mu_1, \sigma_1^2)\mathcal{N}(x; \mu_2, \sigma_2^2) \propto \mathcal{N}\left(x; \frac{\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}, \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}\right)$$

or (eg. EP4 Q6):

$$\begin{aligned} \mathcal{N}(x_n; ax_{n-1}, \sigma^2) &= \mathcal{N}(ax_{n-1}; x_n, \sigma^2) \\ &= \mathcal{N}\left(a, \frac{x_n}{x_{n-1}}, \frac{\sigma^2}{x_{n-1}^2}\right) \end{aligned}$$