Lecture2

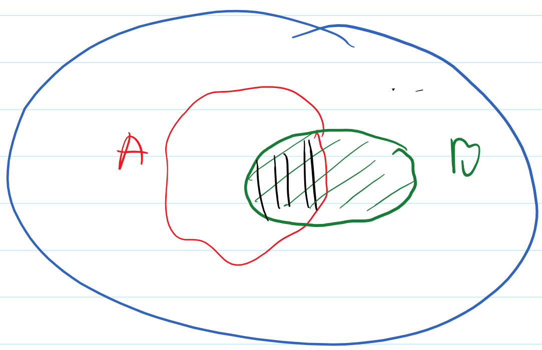
Thursday, September 6, 2018 9:15 AM

LAST TIME:

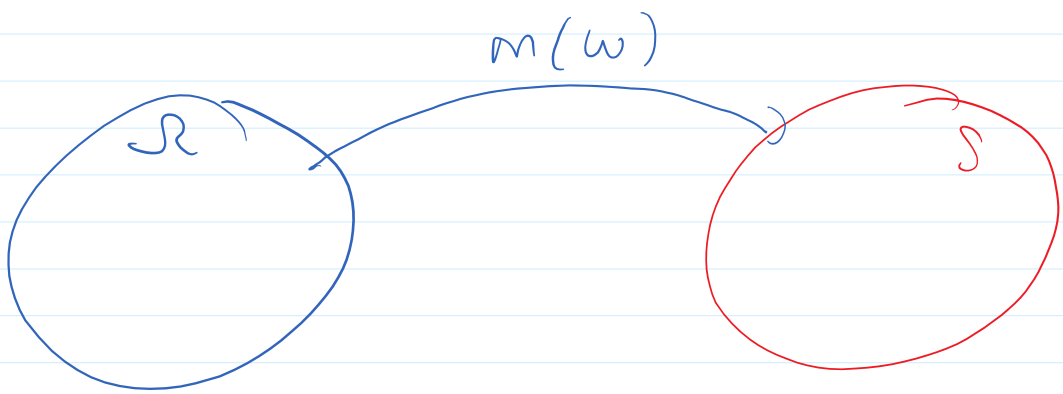
1) Kolmogorov formula

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\text{area}(\cdot)}{\text{area}(\cdot)}$$

area measure



2) Random variable:



distribution of m, denoted as $\mu_m(A)$, which has been shown to be a probability measure

$$\mu_m(A \subset S) = \mathbb{P}[m^{-1}(A) \in \Omega]$$

$$= \mathbb{P}\left(\{ m(\omega) \in A \} \right)$$

* μ_m : is called the pushforward measure of $\mathbb{P}$ via / through $m(\omega)$

$$\mu_m = m_{\#} \mathbb{P}$$

TODAY

Bayes formula

Definition : $\pi_m(m)$, or $\pi(m)$ is called the probability density (w.r.t Lebesgue measure) of a random m if

$$\mu_m(A \subset S) := \int_A \pi(m)\, d\lambda(m)$$

where λ is the standard Lebesgue measure

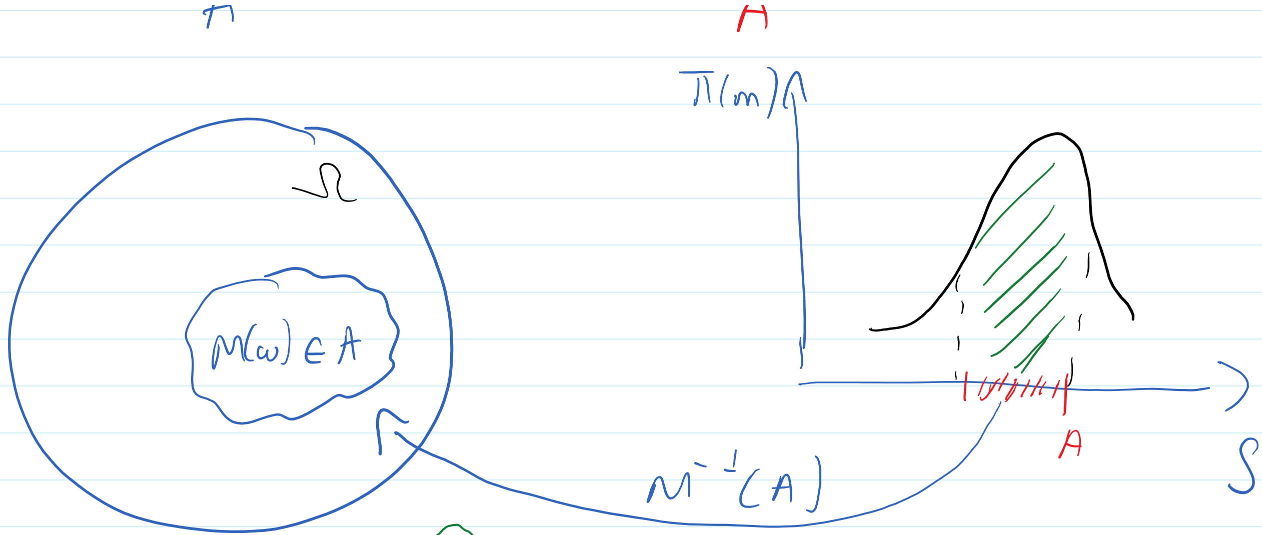
- If the density is known, we don't need the sample space anymore because

$$\mu_m(A) := \mathbb{P}\left(\{ m(\omega) \in A \} \right)$$

$\parallel$ if $\pi(m)$ is known

$$\int_A \pi(m)\, d\lambda(m) = \int_A \pi(m)\, \lambda(dm)$$

$\pi(m)\lambda$



"Area" :  is exactly $\mu_M(A)$

$$\mu_M(A) = \mathbb{P}[\{M(\omega) \in A\}]$$

- again

$$\mu_m(A) = \int_A \pi(m)\,\lambda(dm)$$

formally $\Downarrow$ similar to fundamental theorem of calculus

$$\boxed{\frac{d\mu_m}{d\lambda}(m) = \pi(m)}$$

Side note

$$F(x) - F(0) = \int_0^x f(t)\,dt$$

$$\Uparrow$$

$$\frac{dF}{dx} = F'(x) = f(x)$$

$\frac{d\mu_m}{d\lambda}$ is called the Radon-Nikodym derivative of μ_m w.r.t. λ. One can show that the ordinary derivative is a special case of Radon-Nikodym derivative.

<u>Definition</u>: (mean = expectation)

The expectation or the mean of a random variable $m(\omega)$ is defined as

$$\mathbb{E}[m] := \int_S m\, \pi(m)\, \lambda(dm)$$

$$= \int_S m\, \pi(m)\, dm = \bar{m}$$

<u>Ex 4.1</u>: $\Omega = [-2, 2]$, and define

$$\underline{\mathbb{P}}[B \subset \Omega] = \int_B \frac{1}{4}\, d\Omega = \int_B \frac{1}{4}\, d\lambda$$

Define a random variable $m : \Omega \to S \subset \mathbb{R}$

$$m(\omega) = \begin{cases} 2 & \text{if } \omega \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$; \; 0 \quad \text{otherwise}$

$$\mathbb{E}\{m\} = \int_{S=\{0,2\}} m\,\pi(m)\,dm =$$

$$\| \quad \mu_m(dm) = \pi(m)\,dm$$

$$\int_{S=\{0,2\}} m\,\mu_m(dm) \quad ;$$

$$\bar{m}^{-1}(S) = \Omega \qquad \| \text{ def of } \mu_m : \quad d\Omega = \bar{m}^{-1}(dm)$$

$$\int_{\Omega} m(\omega)\,\mathbb{P}(d\Omega)$$

$$\| \text{ def } \mathbb{P}$$

$$\int_{\Omega} m(\omega)\,\frac{1}{4}\,d\Omega$$

$$\| \text{ def } m(\omega)$$

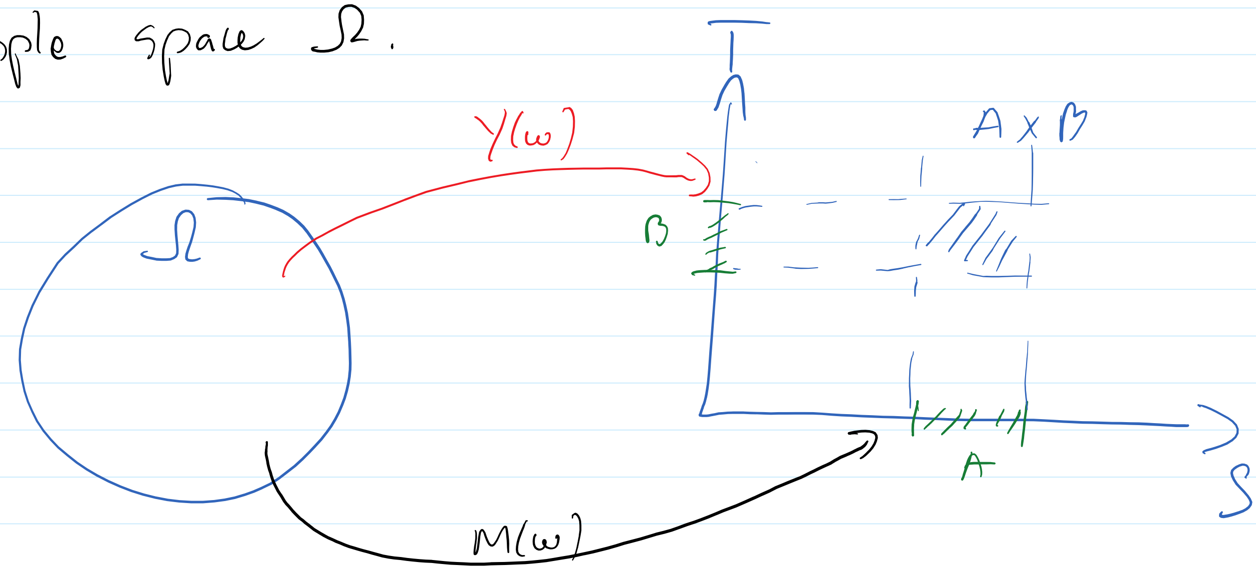
$$\int_0^2 2\,\frac{1}{4}\,d\Omega = 1$$

Def (variance) The variance of $m(\omega)$ is defined as:

$$\mathbb{V}\{m\} := \mathbb{E}\{(m-\bar{m})^2\} = \int (m-\bar{m})^2\,\pi(m)\,dm$$

$$\mathbb{V}[m] := \mathbb{E}[(m-\bar{m})^2] = \int_S (m-\bar{m})^2 \pi(m)\,dm$$

- In Bayesian inversion framework we typically work with joint and conditional probability of many random variables defined on the same sample space Ω.



Def: The joint probability measure of m and y is defined as:

$$\mu_{my}(A \times B) = \mu_{my}(\{m(\omega) \in A\} \cap \{y(\omega) \in B\})$$

|| joint density $\pi_{my}(m,y)$ is known

$$\int \pi_{my}(m,y)\,dm\,dy \qquad \forall A \subset S, B \subset T$$

$A \times B$

Def: m and y are mutually independent if

$$\mu_{my}(A \times B) = \mu_m(A) \times \mu_y(B)$$

Def (marginal density)

- The marginal density of m is the probability of m when y can take any value (irrespective to any value of y):

$$\pi_m(m) = \int_T \pi_{my}(m, y)\, dy$$

Def (Conditional density)

- Conditional density of m given y, $\pi(m|y)$, is defined as

$$\mu(\{m \in A\} | y) = \int_A \pi(m|y)\, dm$$

Thm: (Pre-Bayes' formula).

The conditional density of m given y is given by

$$\pi(m|y) = \frac{\pi(m,y)}{\pi(y)}$$

Proof: the main tool is the kolmogorov formula.

Recall

$$\mu(\{M(\omega)\in A\}|y) := \int_A \pi(m|y)\,dm$$

$\{\omega : M(\omega) \in A\}$

$$\overset{\text{''}}{=}_{\forall A} \int_A \frac{\pi(m,y)}{\pi(y)}\,dm$$

|| def

$$\mathbb{P}[\{M\in A\} \mid Y=y] = \lim_{\Delta y\to 0} \mathbb{P}[\{M\in A\} | y\le Y\le y+\Delta y]$$

|| kolmogorov

$$\lim_{\Delta y\to 0} \frac{\mathbb{P}[\{M\in A\},\ y\le Y\le y+\Delta y]}{\mathbb{P}[y\le Y\le y+\Delta y]}$$

|| Fubini

$$\lim_{\Delta y\to 0} \frac{\int_y^{y+\Delta y}\left[\int_A \pi(m,y)\,dm\right]dy}{\int_y^{y+\Delta y} \cdots}$$

$$\lim_{\Delta y \to 0} \frac{\cdots}{\int_y^{y+\Delta y} \pi(y)\, dy}$$

$$\| $$

$$\lim_{\Delta y \to 0} \frac{\int_A \pi(m,y)\, dm\, \cancel{\Delta y}}{\pi(y)\, \cancel{\Delta y}}$$

$$\| $$

$$\int_A \frac{\pi(m|y)}{\pi(y)}\, dm \qquad \forall\, A \subset S$$

and this ends the proof.

Thm (Bayes' formula).

There holds:

$$\boxed{\pi(m|y) = \frac{\pi(y|m) \times \pi(m)}{\pi(y)}}$$

Proof: from Pre-Bayes we know that

$$\pi(m|y) = \frac{\pi(m,y)}{\pi(y)}$$

$$\pi(y)$$

and $$\pi(y|m) = \frac{\pi(m,y)}{\pi(m)}$$

$$\boxed{\pi(m|y) = \frac{\pi(y|m) \times \pi(m)}{\pi(y)}}$$

<u>Def</u>: $\pi(m)$ is called the prior density (prior knowledge) <u>before</u> we see/conduct experiments or data/observation

<u>Def</u>: $\pi(y|m)$ is called the likelihood of $\underline{m}$
= probability density of y given m. This encodes "information"/knowledge from the data/observation y.

<u>Def</u> $\pi(m|y)$: is called the posterior density. It is the updated knowledge from the prior knowledge when the data y come.

<u>Def</u>: $\pi(y)$ is called the "evidence". It is nothing more than the marginal distribution of y:

$$\pi(y) = \int \pi(y|m) \times \pi(m)\, dm$$

$$\pi(y) = \int_S \pi(y|m) \times \pi(m)\, dm$$

Notation:

$$\pi_{like}(m) = \pi(y|m)$$

$$\pi_{prior}(m) = \pi(m).$$