# Expectation - Back to the Beginning

$\Omega$ finite, $X : \Omega \to \mathbb{R}$, $\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\,\mathbb{P}(\{\omega\}) = \int X \, d\mathbb{P}$

If we also have $g : \mathbb{R} \to \mathbb{R}$, $\mathbb{E}[g(X)] = \sum_{\omega \in \Omega} g(X(\omega))\,\mathbb{P}(\{\omega\})$

$X$ takes finitely many values $t_1, t_2, \dots, t_n$

$$= \sum_{j=1}^{n} \sum_{\substack{\omega \in \Omega : \\ X(\omega) = t_j}} g(X(\omega))\,\mathbb{P}(\{\omega\})$$

$$= \sum_{j=1}^{n} g(t_j) \sum_{\substack{\omega \in \Omega \\ X(\omega) = t_j}} \mathbb{P}(\{\omega\})$$

$$= \sum_{j=1}^{n} g(t_j)\,\mathbb{P}(\{\omega \in \Omega : X(\omega) = t_j\}) = \sum_{j=1}^{n} g(t_j)\,\mathbb{P}(X^{-1}\{t_j\})$$

$$= \sum_{t} g(t)\,\mathbb{P}(X = t)$$

$\mathbb{P} \circ X^{-1} = X_* \mathbb{P} = \mu_X$

$$= \sum_{t} g(t)\,\mu_X(\{t\})$$

$$= \int_{\mathbb{R}} g \, d\mu_X$$

# Change of Variables

**Proposition:** Let $X:(\Omega,\mathcal{F},\mu)\to(S,\mathcal{B})$, $\nu = X_*\mu$.

Then $g:S\to\mathbb{R}$ is in $L^1(S,\mathcal{B},\nu)$
iff $g\circ X:\Omega\to\mathbb{R}$ is in $L^1(\Omega,\mathcal{F},\mu)$, and

$$\int_S g\,d\nu = \int_\Omega (g\circ X)\,d\mu$$

Pf. [HW6]

In particular, take $(S,\mathcal{B}) = (\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$, $\mathbb{X}:(\Omega,\mathcal{F},\mathbb{P})\to\mathbb{R}^d$ a Borel random vector. Then

$\mathbb{X}_*\mathbb{P} = \mathbb{P}\circ\mathbb{X}^{-1}:\mathcal{B}(\mathbb{R}^d)\to[0,1]$ is the distribution/law $\mu_{\mathbb{X}}$.

$$\mathbb{E}[g(\mathbb{X})] = \int g\circ\mathbb{X}\,d\mathbb{P} = \int_{\mathbb{R}^d} g\,d(\mathbb{X}_*\mathbb{P}) = \int_{\mathbb{R}^d} g\,d\mu_{\mathbb{X}}.$$

Radon measure

In particular, for $\mathbb{R}$-valued random variables,

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g \, d\mu_X .$$

If $F_X(t) = \mathbb{P}(X \leq t)$ is the CDF, then $\mu_X = \mu_{F_X}$

and so if $g$ happens to be Riemann-Stieltjes integrable (e.g. continuous),

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(t) \, dF_X(t) = \int_{\mathbb{R}} g(t) \underbrace{F_X'(t)}_{f_X \text{ density of } X.} dt \quad \text{if } F_X \in C^1 .$$

Upshot: if we know $\mu_X$, we can calculate expectations of functions of $X$.

Eg. (moments) $\mathbb{E}[X^n] = \int_{\mathbb{R}} t^n \mu_X(dt)$ ← could be $\infty$.

Note: $t^{n+1} \geq t^n$ for $|t| \geq 1$. → If $\int t^{n+1} \mu_X(dt) < \infty$

$\int_{-1}^{1} \mu_X(dt) < \infty$ $\Rightarrow$ $\int t^n \mu_X(dt) < \infty$.

E.g. $\mathcal{N}(\alpha, t)$ Normal distributions.

$\alpha \in \mathbb{R}$, $t > 0$

Density $\frac{1}{\sqrt{2\pi t}} e^{-(x-\alpha)^2/2t}$

Fact: If $X \overset{d}{=} \mathcal{N}(\alpha, t)$ and $Z \overset{d}{=} \mathcal{N}(0,1)$, then $X \overset{d}{=} \sqrt{t} Z + \alpha$.

$$\mathbb{E}[g(\sqrt{t}Z + \alpha)] = \int g(\sqrt{t}x + \alpha)\mu_Z(dx) = \int g(\sqrt{t}x+\alpha)\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

$$= \int g(u)\frac{1}{\sqrt{2\pi}}e^{-\left(\left(\frac{x-\alpha}{\sqrt{t}}\right)^2\right)/2}\frac{du}{\sqrt{t}}$$

$$= \int g(u)\frac{1}{\sqrt{2\pi t}}e^{-(u-\alpha)^2/2t}du$$

$$= \mathbb{E}[g(X)].$$

$u = \sqrt{t}x + \alpha$

$x = \frac{(u-\alpha)}{\sqrt{t}}$

$dx = \frac{1}{\sqrt{t}}du$

E.g. $g = \mathbb{1}_{(a,b]}$.

$$\mathbb{E}[\mathbb{1}_{(a,b]}(X)] = \mathbb{E}[\mathbb{1}_{(a,b]}(\sqrt{t}Z + \alpha)]$$

$$\begin{array}{ccc} \| & & \| \\ P(X \in (a,b]) & & P(\sqrt{t}Z+\alpha \in (a,b]) \\ \| & & \| \\ \mu_X(a,b] & = & \mu_{\sqrt{t}Z+\alpha}(a,b] \end{array}$$

$\therefore \mu_X = \mu_{\sqrt{t}Z+\alpha}$.

///

Let $Z \overset{d}{=} \mathcal{N}(0,1)$. Then $\mathbb{E}[Z] = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = \lim_{r\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} x e^{-x^2/2}\,dx$

$$\mathbb{E}[Z^2] = \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \lim_{r\to\infty} \int_{-r}^{r} \underset{u}{x} \cdot \underbrace{x e^{-x^2/2}}\,dx \qquad = \lim_{r\to\infty} \frac{1}{\sqrt{2\pi}} \left( -e^{-x^2/2} \right|_{-r}^{r} = 0.$$

$$= 1.$$

$$\therefore \mathbb{E}[\underset{\overset{\parallel d}{\mathcal{N}(\alpha, t)}}{X}] = \mathbb{E}[\sqrt{t} Z + \alpha] = \sqrt{t}\, \underset{0}{\cancel{\mathbb{E}[Z]}} + \mathbb{E}[\alpha] = \alpha.$$

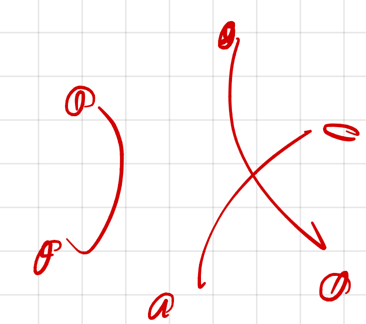
For higher moments, we can use Gaussian integration by parts:

$$\mathbb{E}[Z f(Z)] = \mathbb{E}[f'(Z)]$$

$$\therefore \mathbb{E}[Z^{n+2}] = \mathbb{E}[Z \cdot Z^{n+1}] = (n+1)\,\mathbb{E}[Z^n]$$

$$\hookrightarrow \mathbb{E}[Z^n] = \begin{cases} 0 & \text{if } n \text{ odd} \\ \underbrace{(n-1)(n-3)(n-5)\cdots(3)(1)}_{(n-1)!!} & \text{if } n \text{ is even} \end{cases}$$

$$= \#\text{ pairings of } n \text{ objects}$$



E.g. Exponential distributions $\mathrm{Exp}(\alpha)$, $\alpha > 0$

$T \overset{d}{=} \mathrm{Exp}(\alpha)$ if $F_T(t) = (1 - e^{-\alpha t}) \mathbb{1}_{t \geq 0}$

$\therefore f_T(t) = \frac{d}{dt}(1 - e^{-\alpha t}) = \alpha e^{-\alpha t} \mathbb{1}_{[0,\infty)}$

Often models a "waiting time"

$\mathbb{P}(T > t) = 1 - \mathbb{P}(T \leq t) = 1 - F_T(t) = e^{-\alpha t}$.

$\mathbb{E}[T] = \int_0^\infty \underbrace{t \cdot \alpha e^{-\alpha t}}_{} dt$

$\alpha \cdot t e^{-\alpha t} = -\alpha \frac{\partial}{\partial \alpha} e^{-\alpha t}$

$\downarrow$

$= -\alpha \frac{d}{d\alpha} \int_0^\infty e^{-\alpha t} dt = -\alpha \frac{d}{d\alpha}\left(\frac{1}{\alpha}\right) = \frac{1}{\alpha}$.

E.g. Poisson distribution $\mathrm{Pois}(\alpha)$

$N \overset{d}{=} \mathrm{Pois}(\alpha)$ iff $N : \Omega \to \mathbb{N}$ and

$$P(N=k) = e^{-\alpha} \frac{\alpha^k}{k!} = \mu_N(\{k\})$$

Models number of events occurring in a fixed time period.

Closely related to $\mathrm{Exp}(\alpha)$; we'll look into this later.

$$\mathbb{E}[N] = \int_{\mathbb{N}} t \cdot \mu_N(dt) = \sum_{k=0}^{\infty} k \cdot \mu_N(\{k\})$$

$$= \sum_{k=0}^{\infty} k e^{-\alpha} \frac{\alpha^k}{k!} \quad \Leftarrow \quad k\alpha^k = \alpha \frac{\partial}{\partial \alpha} \alpha^k$$

$$= e^{-\alpha} \sum_{k=0}^{\infty} \alpha \frac{\partial}{\partial \alpha} \frac{\alpha^k}{k!} = \alpha e^{-\alpha} \frac{d}{d\alpha} \underbrace{\sum_{k=0}^{\infty} \frac{\alpha^k}{k!}}_{e^{\alpha}} = \alpha.$$

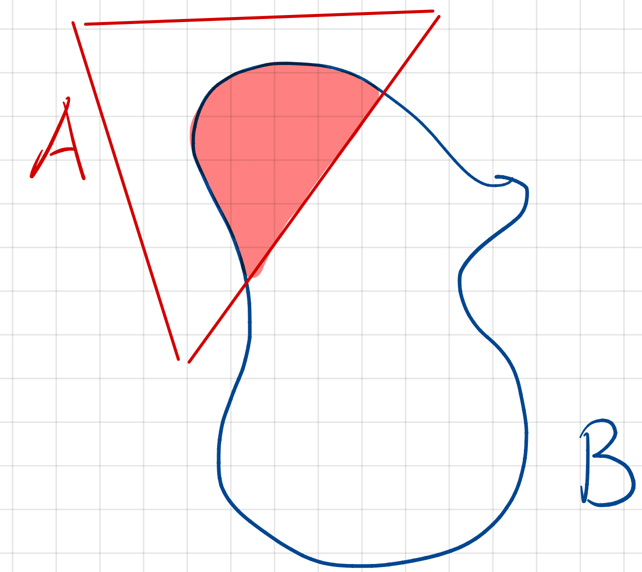
Eg. Uniform Random Variables / Vectors

Let $B \in \mathcal{B}(\mathbb{R}^d)$ with $\lambda(B) > 0$.

A random vector $\underline{X} \overset{d}{=} \mathrm{Unif}(B)$ iff $\mathbb{P}(\underline{X} \in A) = \dfrac{\lambda(A \cap B)}{\lambda(B)}$

$$\mathbb{P}(\underline{X} \in A) = \frac{1}{\lambda(B)} \lambda(A \cap B)$$

$$= \frac{1}{\lambda(B)} \int_A \mathbb{1}_B \, d\lambda$$




$\therefore \; f_{\underline{X}} = \frac{1}{\lambda(B)} \mathbb{1}_B$. wrt $\lambda$ on $\mathbb{R}^d$

For example: $U \overset{d}{=} \mathrm{Unif}([a,b])$

$$f_U(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}$$

$$E[U] = \int_{\mathbb{R}} x \cdot \frac{1}{b-a} \mathbb{1}_{[a,b]}(x) \, dx = \frac{1}{b-a} \int_a^b x \, dx = \frac{1}{b-a} \frac{(b^2 - a^2)}{2} = \frac{a+b}{2}$$