Measure Theory

$(\Omega, \mathcal{F}, \mathbb{P})$,

$\mathbb{P}: \mathcal{F} \to S$, usually $S=[0,1]$

$\mathcal{F} \subseteq \mathcal{P}(\Omega)$, observable

Measure properties:

- $\mathbb{P}[\Omega]=1$
- σ-addictivity: $\mathbb{P}[\bigcup_{k}^{\infty} A_k] = \sum^{\infty} \mathbb{P}[A_k]$ countably addictive!
- F is σ-field:

$\Omega, \emptyset \in \mathcal{F}$; closed under \ and countable ∪

consequences

$A\subseteq B \Rightarrow \mathbb{P}[B\setminus A] = \mathbb{P}[B]-\mathbb{P}[A]$

$A\subseteq B \Rightarrow \mathbb{P}[A] \le \mathbb{P}[B]$ so monotone incr

$\mathbb{P}[A\cup B] = \mathbb{P}[A]+\mathbb{P}[B]-\mathbb{P}[A\cap B]$

If $A_1 \subseteq A_2 \subseteq \dots \subseteq \Omega$, $\mathbb{P}[\bigcup A_k] = \lim_{k\to\infty} \mathbb{P}[A_k]$ "monotone cont."

Fact σ-addictivity ⇔ monotone cont.

Generating σ-field — infinite coin flip

$\mathcal{F} = \sigma(\{\{x_1,\dots,x_n\} \in \{0,1\}^n \mid n \ge 1\})$

Descrete Model

$\sum_{w\in\Omega} \mathbb{P}[\{w\}] = 1$

$p_k := \mathbb{P}[\{w_k\}]$

For unif. disjoint finite, $\mathbb{P}[A] = \frac{|A|}{|\Omega|}$

Combinatorics

Choose size k subset from n:

$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} = \binom{n}{k,\,(n-k)}$ ← correction

Choosing multiple distinguishable subsets

$\binom{n}{k_1,k_2} = \binom{n}{k_1}\binom{n-k_1}{k_2}$

$\binom{n}{k_1,k_2,\dots,k_m} = \frac{k}{k_1!\,k_2!\dots k_m!}$ if $\sum k_i = n$

For indistinguishable partitions, correction!

7 into 4-1-1-1 : $\binom{7}{4\ \underbrace{1\ 1\ 1}}\frac{1}{3!}$

RVs

$X: \Omega \to S$ usually $\mathbb{R}$ discrete if Im(X) countable

↑ the input is random

Let $\mathcal{G}$ be a σ-field on S, $G \in \mathcal{G}$

$\mathbb{P}[X\in G] = \mathbb{P}[\{w \mid X(w) \in G\}]$
$= \mathbb{P}[\{w \mid w \in X^{-1}(G)\}]$
$= \mathbb{P}[X^{-1}(G)]$
$= \mathbb{P}\circ X^{-1}(G)$
$= \mu_X(G)$

← measure as in $(S, \mathcal{G}, \mu_X)$ and distribution of X in $(\Omega, \mathcal{F}, \mathbb{P})$

Correlation

$\mathrm{corr}(X,Y) := \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}X}\sqrt{\mathrm{var}Y}} \in [-1,1]$

$\mathrm{corr}(X,X) = \frac{\mathrm{cov}(X,X)}{\mathrm{var}X} = 1$

Moments

k-th moment is $\mathbb{E}[|X|^k]$

$\mathbb{E}[|X|^k]^{\frac{1}{k}} = \|X\|_k$ is k-th norm

$\mathcal{L}^p = \{X \mid \|X\|_p \text{ finite}\}$

Thm if $1 \le q < p$ then $\mathcal{L}^p \subseteq \mathcal{L}^q$

viz. $\mathcal{L}^2 \subseteq \mathcal{L}^{1.5} \subseteq \mathcal{L}^1$ etc.

Use $\varphi(x) = |x|^{p/q}$, $Y = |X|^q$

$\mathbb{E}[(|X|^q)^{p/q}] \ge (\mathbb{E}[|X|^q])^{p/q}$

$\|X\|_p \ge \|X\|_q$

Independence

$A \perp B \Leftrightarrow \mathbb{P}[A\cap B] = \mathbb{P}[A]\mathbb{P}[B]$
$\Leftrightarrow \mathbb{P}[A|B] = \mathbb{P}[A]$

ℝ RVs $X \perp Y \Leftrightarrow \forall A,B \subseteq \mathbb{R}$, $\{X\in A\} \perp \{Y\in B\}$

more generally $\mathbb{P}[\bigcap X_k \in A_k] = \prod \mathbb{P}[X_k \in A_k]$

$X \perp Y$:
- $\Rightarrow \mathbb{E}XY = \mathbb{E}X\cdot\mathbb{E}Y$
- $\Rightarrow \mathrm{cov}(X,Y)=0$
- $\Leftrightarrow \mu_X\mu_Y(x,y) = \mu_{X,Y}(x,y)$
- $\Leftrightarrow F_{X,Y}(x,y) = F_X(x)F_Y(y)$
- $\Leftrightarrow f_{X,Y}(x,y) = f_X(x)F_Y(y)$
- $\Rightarrow \mathrm{corr}(X,Y)=0$

$X_1,\dots,X_n$ indep $\overset{\not\Leftarrow}{\Rightarrow}$ X_i, X_j pairwise indep

Expectation

Discrete: $\mathbb{E}X = \sum_{x\in \mathrm{Im}X} x\,\mathbb{P}[X=x]$
$\equiv \min_{b\in\mathbb{R}} (X-b)^2$
$\equiv \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k$ X_k iid ~X

- Not every RV have $\mathbb{E}$
- If X is indicator, $\mathbb{E}X = \mathbb{P}[X=1]$
- $\mathbb{E}[\cdot]$ is linear
- $(\forall w, X(w) \ge Y(w)) \Rightarrow \mathbb{E}X \ge \mathbb{E}Y$
- $\mathbb{E}[\cdot]$ is monotone cont: let RVs $0 \le X_1 \le \dots$ suppose $X_n \to X$, then $\mathbb{E}X = \lim_{n\to\infty}\mathbb{E}[X_n]$

If $X\in\mathbb{N}$, $\mathbb{E}X = \sum_{k=0}^{\infty}\mathbb{P}[X>k]$

$\mathbb{E}X = \mathrm{sum}\begin{pmatrix} p_1 & p_2 & p_3 \\ & p_2 & p_3 & \cdots \\ & & p_3 \end{pmatrix} = \mathrm{sum}\begin{pmatrix}\mathbb{P}[X>0] \\ \mathbb{P}[X>1] \\ \vdots\end{pmatrix}$

Partition: $\mathbb{E}X = \sum \mathbb{E}[X|B_k]\mathbb{P}[B_k]$ where $\Omega = \bigcup B_k$

$\mathbb{E}[g(X)] = \sum_{x\in\mathrm{Im}X} g(x)\mathbb{P}[X=x]$

If $X \ge 0$, $\mathbb{E}X = \int_0^\infty \mathbb{P}[X>t]\,dt$

$\mathbb{E}\,\Phi(X) = \int_0^\infty \mathbb{P}[X>t]\Phi(t)\,dt$

Variance

$\mathrm{var}X = \sigma^2 = \mathbb{E}[(X-\mathbb{E}X)^2]$
$= \mathbb{E}[X^2] - (\mathbb{E}X)^2$

$\mathrm{var}\,\alpha X = \alpha^2\,\mathrm{var}X$

$\mathrm{var}(X+c) = \mathrm{var}X$

$\mathrm{var}X = \mathrm{cov}(X,X)$

$\mathrm{var}(X+Y) = \mathrm{var}X + \mathrm{var}Y + 2\underbrace{\mathrm{cov}(X,Y)}_{0 \text{ if } X,Y \text{ uncorrelated}}$ ← cov(X+Y, X+Y) …

more generally:

$\mathrm{var}(\sum_{k=1}^n X_k) = \sum_{k=1}^n \mathrm{var}X_k + 2\sum_{i<j}\mathrm{cov}(X_i,X_j)$

$\mathrm{cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]$
$= \mathbb{E}[XY] - \mathbb{E}X\cdot\mathbb{E}Y$

↑ bilinear, symmetric

$X\perp Y \overset{\not\Leftarrow}{\Rightarrow} \mathrm{cov}(X,Y) = \mathbb{E}X\cdot\mathbb{E}Y - \mathbb{E}X\cdot\mathbb{E}Y = 0$

corollary: if $X_1,\dots,X_n$ indep, $\mathrm{var}(\sum X_k) = \sum \mathrm{var}X_k$

$\mathrm{cov}(\alpha X, Y) = \alpha\,\mathrm{cov}(X,Y)$

$\mathrm{cov}(X+Z,Y) = \mathrm{cov}(X,Y) + \mathrm{cov}(Z,Y)$

$\mathrm{cov}(X+b,Y) = \mathrm{cov}(X,Y)$

Conditional

Condition on B: $\Omega' = B$
$\mathcal{F}' = \{A\cap B \mid A\in\mathcal{F}\}$
$\mathbb{P}_B(A') = \frac{\mathbb{P}(A')}{\mathbb{P}(\Omega')} = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}$

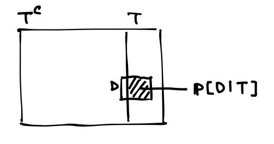
$\mathbb{P}_B[\cdot] = \mathbb{P}[\cdot|B]$ by $\mathbb{P}[A|B] = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}$

$\mathbb{P}[A|B] = \mathbb{P}[B|A]\frac{\mathbb{P}[A]}{\mathbb{P}[B]}$

Partitioning $\mathbb{P}[A] = \sum\mathbb{P}[A\cap B_k]$ with $\Omega = \bigcup^{N\le\infty} B_k$
$= \sum\mathbb{P}[A|B]\mathbb{P}[B]$

In reverse, $\mathbb{P}[B_j|A] = \mathbb{P}[A|B_j]\frac{\mathbb{P}[B_j]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A|B_j]\mathbb{P}[B_j]}{\sum\mathbb{P}[A|B_k]\mathbb{P}[B_k]}$

Disease test intuition



Multiple tests assumed indep w.r.t. $\mathbb{P}[\cdot|D]$

$\mathrm{var}_A(X) = \mathrm{var}(X|A) = \mathbb{E}_A[(X-\mathbb{E}_AX)^2]$

$\mathbb{P}[X\in B|Y=s] = \int_B f_{X|Y}(x,s)\,dx$

$\mathbb{E}[g(X,Y)|Y=s] = \int_{\mathbb{R}} g(x,s)f_{X|Y}(x,s)\,dx$

$f_{X|Y}(x,y) := \begin{cases}\frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{if } f_Y(y)>0 \\ 0 & \text{else}\end{cases}$

$\mathbb{P}[X\in B] = \int_{\mathbb{R}} f_Y(y)\,\mathbb{P}[X\in B|Y=y]\,dy$

$\int_{\mathbb{R}} f_Y(y)\int_B f_{X|Y}(x,y)\,dx\,dy$

$\int_B\int_{\mathbb{R}} f_Y(y) f_{X|Y}(x,y)\,dy\,dx$

$\int_B\int_{\mathbb{R}} f_{X,Y}(x,y)\,dy\,dx$

Continuous Model

X abs. cont. ⇔ ∃ pdf $f_X: \mathbb{R}\to\mathbb{R}^+$
1. $f_X \ge 0$
2. $\int_{\mathbb{R}} f_X(x)\,dx = 1$
3. $\forall B \subseteq \mathbb{R}$, $\mathbb{P}[X\in B] = \int_B f_X(x)\,dx$ ← usually assume $B = (a,b] \subseteq \mathbb{R}$ in proof

Write $\int f_X(x)\,dx = \int \mu_X(dx)$

- $\mathbb{E}[g(X)] = \int g(x)\,\mu_X(dx)$
- $\mathrm{var}X = \int x^2\mu_X(dx) - (\mathbb{E}X)^2$

CDF F_X (F_X and μ_X uniquely determine e.o.)

$F_X(t) = \mathbb{P}[X\le t]$
$= \mu((-\infty, t])$

1. F_X monotone increasing
2. $\lim_{t\to-\infty}F_X(t) = 0$, $\lim_{t\to\infty}F_X(t) = 1$
3. Right continuous

$\frac{d}{dt}\int_{-\infty}^t f_X(x)\,dx \overset{FTC}{=} f_X(t)$

X abs. cont. $\overset{\not\Leftarrow}{\Rightarrow}$ X cont. ($\mathbb{P}[X\in B] = 0$ ∀ countable B)

$\vec{X} = (X,Y)$ abs cont if ∃ pdf $f_{X,Y}$

$f_{X,Y}(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}\mathbb{P}[X\le x, Y\le y]$
$= \frac{\partial}{\partial x}\frac{\partial}{\partial y}F_{X,Y}(x,y)$
$= \frac{\partial}{\partial x}\frac{\partial}{\partial y}\int_{-\infty}^{x}\int_{-\infty}^{x} f_{X,Y}(s,t)\,dt\,ds$

Multivar Normal

$\vec{X} = (X_1,\dots,X_n)$, X_i iid ~ N(0,1)

$\vec{Y} = A\vec{X}$ ← linear transformed

$\to \Phi(\vec{x}) = A\vec{x}$, $[D\Phi]_{\vec{x}} = A$ $\forall\vec{x}$ since Φ linear

$f_{\vec{Y}}(\vec{y}) = f_{\vec{X}}(\vec{x})\,|\det[D\Phi]_{\vec{x}}|^{-1}$
$= f_{\vec{X}}(\Phi^{-1}(\vec{y}))\frac{1}{|\det A|}$
$\dots = \frac{1}{|\det A|}\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}\vec{y}^T(AA^T)^{-1}\vec{y}}$
$\dots = \frac{1}{\sqrt{\det C\cdot(2\pi)^n}}e^{-\frac{1}{2}\vec{y}^TC^{-1}\vec{y}}$ fact $\sqrt{\det C} = |\det A|$

$C = AA^T$ is the covariance matrix ← symmetric, positive

$\mathrm{cov}(y_k, y_j) = (AA^T)_{ij} = C_{ij}$

Correlated joint normal:

$X,Y \sim N(0,1)$, $C = \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}$ so $\mathrm{corr}(X,Y) = \rho$ ← write $(X,Y)\sim N(0,1,\rho)$

$Y|_{X=x} \sim N(\rho x, 1-\rho^2)$

$f_{X,Y}(x,y) \overset{\triangle}{=} \frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2}\frac{1}{1-\rho^2}(x^2-2\rho xy+y^2)}$

With shift $f_{\vec{Y}}(y) = \frac{1}{\sqrt{\det C\,(2\pi)^n}}e^{-\frac{1}{2}(\vec{y}-\vec{b})^TC^{-1}(\vec{y}-\vec{b})}$

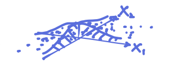
$\mathbb{E}\vec{Y} = \vec{b}$

General $X\sim N(\mu_X,\sigma_X^2)$ $Y\sim N(\mu_Y,\sigma_Y^2)$, $\mathrm{corr}(X,Y)=\rho$

$f_{X,Y} = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right)}$

$C = \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}$

⚠ $X_1,\dots,X_n$ normal $\not\Rightarrow$ $X_1,\dots,X_n$ joint normal



Law of small numbers aka rare events

$X_{1..n}$ iid ~ $\mathrm{Ber}(\frac{\lambda}{n})$ $\lambda>0$

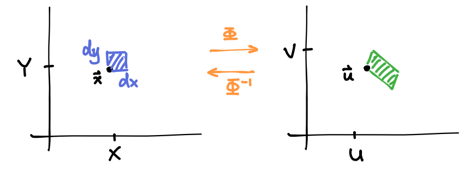
$S_n = \sum X_k \sim \mathrm{Binom}(n, \frac{\lambda}{n})$, $\mathbb{E}S_n = \lambda$ ← rare event

$S_n \xrightarrow[n\to\infty]{} \mathrm{poi}(\lambda)$ ← poission approx. of binom

Transformation

$\Phi(X,Y) = (U,V)$ ← bijective & differentiable

$\Phi\begin{bmatrix}X\\Y\end{bmatrix} = \begin{bmatrix}U(X,Y)\\V(X,Y)\end{bmatrix}$ $\Phi^{-1}\begin{bmatrix}U\\V\end{bmatrix} = \begin{bmatrix}X(U,V)\\Y(U,V)\end{bmatrix}$



$D\Phi = \frac{\partial(u,v)}{\partial(x,y)} = \begin{bmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{bmatrix}$

$D\Phi^{-1} = \frac{\partial(x,y)}{\partial(u,v)} = \begin{bmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{bmatrix}$

$f_{U,V}(\vec{u}) = f_{X,Y}(\vec{x})\cdot|[\det D\Phi]_{\vec{x}}|^{-1}$
$= f_{X,Y}(\vec{x})\cdot|[\det D\Phi^{-1}]_{\vec{u}}|$ where $\vec{x} = \Phi^{-1}(\vec{u})$

Ex. X, Y iid ~ N(0,1)

$\Phi(x,y) = (\overset{R}{\sqrt{x^2+y^2}}, \overset{\theta}{\tan^{-1}(y/x)})$

$\Phi^{-1}(r,\theta) = (r\cos\theta, r\sin\theta)$

$\det D\Phi^{-1} = \det\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & -r\cos\theta\end{bmatrix} = r$

$f_{R,\theta}(r,\theta) = f_{X,Y}(x(\theta,r), y(\theta,r))\cdot|\det D\Phi^{-1}|$
$\dots = \frac{r}{2\pi}e^{-\frac{1}{2}r^2}1_{\mathbb{R}^+}(r)\,1_{[0,2\pi]}(\theta)$

can also show $R\perp\Theta$

Joint dist

marginal recoverable from joint, not always other way

$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy$

Joint: $\vec{X}$ abs. cont. ⇔ ∃ pdf $f_{\vec{X}}: \vec{X}\to\mathbb{R}$,

$\mathbb{P}[\vec{X}\in B] = \int_B f_{\vec{X}}(\vec{x})\,d\vec{x}$
$= \int_{\mathbb{R}}\cdots\int_{\mathbb{R}} f_{X_1,\dots,X_n}(x_1,\dots,x_n)\,dx_1,\dots,dx_n$

$\mathbb{E}[\varphi(\vec{X})] = \int_{\mathbb{R}^n}\varphi(\vec{x})f_{\vec{X}}(\vec{x})\,d\vec{x}$

$\mu_{X+Y}(z) = \sum_{x\in\mathrm{Im}X}\mu_{X,Y}(x, z-y)$
$\overset{\text{if } X\perp Y}{=} \sum_{x\in\mathrm{Im}X}\mu_X(x)\mu_Y(z-y)$ "convolution"

$f_{X+Y}(z) = \int_{\mathbb{R}} f_{X,Y}(x, z-x)\,dx$
$\overset{\text{if } X\perp Y}{=} \int_{\mathbb{R}} f_X(x)f_Y(z-x)\,dx$ — "convolution" $= (f_X * f_Y)(z)$

Partition $\mathbb{E}g(X) = \int f_Y(y)\,\mathbb{E}[g(X)|Y=y]\,dy$

Central Limit

$(X_k)_{k\ge1}$ iid with finite $\mathbb{E}=m$ and $\mathrm{var}=\sigma^2$

$S_n = \sum_{k=1}^n X_k$ Know $\mathbb{E}S_n = nm$, $\mathrm{var}S_n = n\sigma^2$

Thm $\tilde{S}_n := \frac{S_n - \mathbb{E}S_n}{\sqrt{\mathrm{var}S_n}} \sim N(0,1)$ as $n\to\infty$

Ex. Ber(0.6), n = 100

$\mathbb{P}[S_n > 65] = \mathbb{P}[\frac{S_n-\mathbb{E}S_n}{\sqrt{\mathrm{var}S_n}} > \frac{65-\mathbb{E}S_n}{\sqrt{\mathrm{var}S_n}}]$
$= \mathbb{P}[N > \frac{65-60}{\approx 1.02}]$ where $N\sim N(0,1)$

Error bound: let $\rho = \mathbb{E}[|X-\mathbb{E}X|^3]$ be finite

Thm $\mathbb{P}[\tilde{S}_n \le t] = \Phi(t) \pm \frac{3}{\sqrt{n}}\frac{\rho}{\sigma^3}$

Calculus

$\int u\,dv = uv - \int v\,du$

$\int x^2e^x\,dx$ $u = x^2$ $dv = e^x$

u	dv
x^2	e^x
$2x$	e^x
2	e^x
0	e^x

$= x^2e^x - 2xe^x + 2e^x + C$

Fubini: order of iterative integrals doesn't matter

$\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}$

$\frac{d}{dt}\int_{\varphi(t)}^a f_X(x)\,dx = -F'(\varphi(t))\,\varphi'(t)$

$\int_{\mathbb{R}^+} t^k\alpha e^{-\alpha t}\,dt = \frac{k!}{\alpha^k}$

$\int_{\mathbb{R}^+} t^ke^{-t}\,dt = k!$

$f_{T_n}(s) = \frac{\lambda^ns^{n-1}}{(n-1)!}e^{-\lambda s}1_{\mathbb{R}^+}(s)$ ← sum of exp(λ)

$\int \alpha e^{-\alpha t}\,dt = -e^{-\alpha t} + C$

$[-e^{-\alpha t}]_0^\infty = 0 - -1 = 1$

	$\frac{d}{dx}$
$\sin^{-1}x$	$\frac{1}{\sqrt{1-x^2}}$
$\tan x$	$\sec^2x$
$\cot x$	$-\csc^2x$
$\sec x$	$\sec x\tan x$
$\cot x$	$-\csc x\cot x$

Matrix

$A = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$ $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$

Taylor expansion

$e^\lambda = \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$

Geom seq derivative

$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$

Geom series

$S_n = ar^0 + ar^1 + \dots + ar^{n-1}$
$= a\left(\frac{1-r^n}{1-r}\right)$

$S_\infty = \frac{a}{1-r}$

Harmonic

$H_n = \sum_{i=1}^n\frac{1}{i} \le \ln n + 1$

Exp

$\lim_{n\to\infty}(1+\frac{z}{n})^n = e^z$

Arith

$\sum_{i=1}^n i = \frac{n(n+1)}{2}$

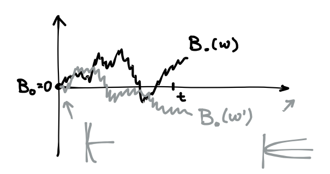
Misc

$(u+1)^n = \sum_{k=0}^n u^k\binom{n}{k}$

$\sum_{j=0}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-1-j} = 1$

$\sum_{n=0}^{\infty}\frac{1}{n!} = e$

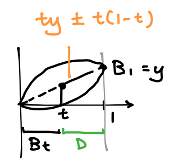
Brownian Motion



BM def:
1. $B_0 = 0$
2. $\forall\, 0<t_1<\dots<t_n$
$D_i := B_{t_i} - B_{t_{i-1}}$ indep
$D_i \sim N(0, t_i - t_{i-1})$ — variance linear to time, std scales with $\sqrt{\text{time}}$
3. $\forall w$, $B_\cdot(w): t \mapsto B_t(w)$ continuous

$f_{B_s|B_t} = f_{B_t|B_s}\frac{f_{B_s}}{f_{B_t}}$

▷ Brownian bridge:



$f_{B_t|B_1=y} \sim N(ty, t(1-t))$

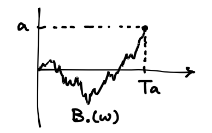
$B_1 = B_t + D$

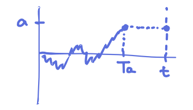
$B_t \perp D$

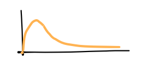
$f_{B_t|B_1=y}(x) = \frac{f_{B_1|B_t}(x,y)\,f_{B_t}(x)}{f_{B_1}(y)}$
$= \frac{\varphi_{1-t}(y-x)\,\varphi_t(x)}{\varphi_1(y)}$

▷ First hitting time

$f_{T_a}(t) = \frac{a}{\sqrt{2\pi t^3}}e^{-\frac{1}{2}\frac{a^2}{t}}1_{(0,\infty]}(t)$







$\mathbb{P}[B_t > a] = \mathbb{P}[B_t > a, T_a \le t]$
$= \mathbb{P}[B_t > a | T_a \le t]\,\mathbb{P}[T_a \le t]$
$= \frac{1}{2}\mathbb{P}[T_a \le t]$ symmetry

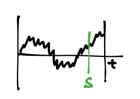
$\mathbb{P}[T_a \le t] = 2\,\mathbb{P}[B_t > a]$

$\frac{d}{dt}\mathbb{P}[T_a\le t] = 2\frac{d}{dt}\int_a^\infty\varphi_t(x)\,dx$
$= 2\frac{d}{dt}(1 - \Phi(\frac{a}{\sqrt{t}}))$
$= f_{T_a}(t)$

▷ Running max $M_s(w) := \max_{t\le s} B_t(w) \sim |B_s|$

$\mathbb{P}[M_s \ge a] = \mathbb{P}[T_a \le s] = \mathbb{P}[|B_s| \ge a]$

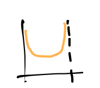
▷ Last visit zero $L_t(w) := \{s\le t \mid B_s = 0\}$



N := num visits in (s,t]

$\mathbb{P}[L_t \le s] = \mathbb{P}[N=0]$
$= \int_{\mathbb{R}} f_{B_s}(x)\,\mathbb{P}[N=0|B_s=x]\,dx$
$= \int_{\mathbb{R}} f_{B_s}(x)\,\mathbb{P}[T_0 > t-s]\,dx$
$\dots = \frac{2}{\pi}\sin^{-1}\sqrt{\frac{s}{t}}$

$f_{L_t}(s) = \frac{1}{\pi\sqrt{s(t-s)}}1_{[0,t]}(s)$

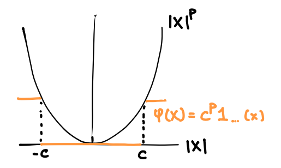


Chebyshev

Fix $p>0$, $c>0$, $\lambda>0$

$\mathbb{P}[|X|\ge c] \le \frac{\mathbb{E}[|X|^p]}{c^p}$

$\mathbb{P}[X\ge c] \le \frac{1}{e^{\lambda c}}\mathbb{E}[e^{\lambda X}]$



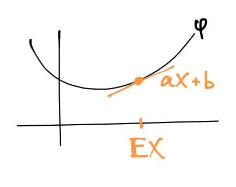
$\varphi(X) \le |X|^p$ ∀w

$\mathbb{E}\,\varphi(X) \le \mathbb{E}\,|X|^p$

Jensen

Let φ(x) be convex, X with finite $\mathbb{E}X$

Then $\mathbb{E}\,\varphi(X) \ge \varphi(\mathbb{E}X)$



$\varphi(X) \ge aX + b$

$\mathbb{E}\,\varphi(X) \ge a\mathbb{E}X + b$
$= \varphi(\mathbb{E}X)$

Distributions

▷ $X\sim\mathrm{Ber}(p)$ $X\in\{0,1\}$

$\mathbb{P}[X=1] = p$ $\mathbb{E}X = p$

$\mathbb{P}[X=0] = 1-p =: q$ $\mathrm{var}X = pq = p(1-p)$

▷ Binom

$X_1,\dots,X_n$ iid ~ Ber(p) $\mathbb{E}X = np$

$X = S_n = \sum^n X_k \sim \mathrm{Binom}(n,p)$ $\mathrm{var}X = np(1-p)$

$\mathbb{P}[X=k] = \mu_X(\{k\}) = \binom{n}{k}p^k(1-p)^{n-k}$

▷ Geom

$X_1, X_2, \dots$ iid ~ Ber(p) $\mathbb{E}X = \frac{1}{p}$

X = time until first success $\mathrm{var}X = \frac{1-p}{p^2}$

$= \min\{k\ge1 \mid X_k = 1\} \in \mathbb{N}^+$

$\mu_X(\{k\}) = \mathbb{P}[X=k] = (1-p)^{k-1}p$

▷ Neg Binom

$X_1, X_2, \dots$ iid ~ Ber(p) $\mathbb{E}X_r = \frac{r(1-p)}{p}$

X_r := time until r-th success $\mathrm{var}X_r = \frac{r(1-p)}{p^2}$

$\mathbb{P}[X_r = k] = p^r(1-p)^{k-r}\binom{k-1}{r-1}$

▷ $X\sim\mathrm{Poi}(\lambda)$ $X\in\mathbb{N}^+$ $\mathbb{E}X = \lambda$

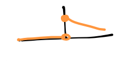
$\mathbb{P}[X=k] = e^{-\lambda}\frac{\lambda^k}{k!}$ $\mathrm{var}X = \lambda$

$\mathbb{E}X^2 = \lambda^2+\lambda$

▷ $X\sim\mathrm{Unif}[a,b]$ $\mathbb{E}X = \frac{b-a}{2}$

$f_X(x) = \frac{1}{b-a}1_{[a,b]}(x)$ $\mathrm{var}X = \frac{1}{12}(b-a)^2$

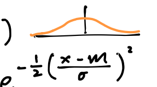
▷ $X\sim\exp(\lambda)$



$\mathbb{E}X = \frac{1}{\lambda}$

$f_X(x) = \lambda e^{-\lambda x}1_{\mathbb{R}^+}(x)$ $\mathrm{var}X = \frac{1}{\lambda}$

▷ $X\sim N(m,\sigma^2)$



$\mathbb{E}X = m$

$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$ $\mathrm{var}X = \sigma^2$

$\mathbb{E}X^2 = \sigma^2 + m^2$

write $\varphi_t(x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{1}{2}\frac{x^2}{t}}$ (t: var)

Fact $\underset{\mathrm{Poi}(\lambda)}{X} + \underset{\mathrm{Poi}(\mu)}{Y} \sim \mathrm{Poi}(\lambda+\mu)$

$X\perp Y$

Card

Prob of 2-1-1-0 Ace dist

choose player* to get 2 $\binom{4}{1}$ 4

0 $\binom{3}{1}$ 3

choose which 2 to give* $\binom{4}{2}$ 6

permute remaining to 1-1 2! 2

144

for each arrangement, choose positions for the Aces

$\binom{13}{2}\cdot2\cdot\binom{13}{1}\binom{13}{1}$ ← permute! . Then permute rest 48!

$144\cdot\binom{13}{2}\cdot2\cdot13^2\cdot48!\,\frac{1}{52!}$ ✓

Disease

$\mathbb{P}\,T|D = 99\%$

$\mathbb{P}\,T^c|D^c = 97\%$

$\mathbb{P}\,D|T = \mathbb{P}\,T|D\,\frac{\mathbb{P}D}{\mathbb{P}T}$
$= \frac{\mathbb{P}(T|D)\ \mathbb{P}(D)}{\mathbb{P}(T|D)\,\mathbb{P}(D) + \mathbb{P}(T|D^c)\,\mathbb{P}(D^c)}$
$= \frac{0.99\ \ 0.001}{0.99(0.001) + (0.03)(0.999)}$
≈ 0.0319
$\approx \frac{1}{31}$

Correlated Joint Normal

X, Y ~ N(0,1)

let X, Z ~ iid N(0,1)

$Y := \alpha X + \beta Z$. Want $C = \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}$

$\mathrm{cov}(X,Y) = \mathrm{cov}(X, \alpha X + \beta Z)$
$= \mathrm{cov}(X,\alpha X) + \mathrm{cov}(X,\beta Z)$
$= \alpha = \rho$

$\mathrm{var}\,Y = \mathrm{var}(\rho X + \beta Z)$
$1 = \rho^2 + \beta^2$
$\beta = \sqrt{1-\rho^2}$

$\Rightarrow Y = \rho X + \sqrt{1-\rho^2}\,Z$

Rand num of coin flips

N ~ poi(λ) N flips Ber(p)

X := num heads Y := num tails

Fact X⊥Y. X ~ poi(pλ)

Given X = k, Y ~ poi((1−p)λ) still

n-dim simplex with side len a

$\sigma_n(a) = \frac{a^n}{n!}$

Cond exp & var

$\mathbb{E}[\mathbb{E}[X|Y]] = \int\mathbb{E}[X|Y=y]\,f_Y(y)\,dy = \mathbb{E}X$

$\mathrm{var}X = \mathbb{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbb{E}[X|Y])$

$X_{1..n}$ iid ~ Unif [0,a] $Z_n = \max(X_1,\dots,X_n)$

$U_n = n(a - Z_n)$

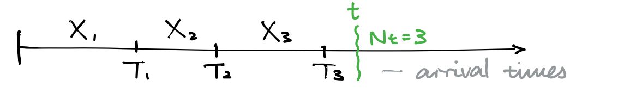
WTS $U_n \xrightarrow[n\to\infty]{} \exp(\lambda)$ for some λ

$\mathbb{P}[U_n \le t] = \mathbb{P}[n(a-Z_n) \le t]$
$= 1 - \mathbb{P}[Z_n \le a - \frac{t}{n}]$
$= 1 - \left(1 + \frac{-\frac{t}{a}}{n}\right)^n$

$\mathbb{P}[U_n \le t] \xrightarrow[n\to\infty]{} 1 - e^{-\boxed{\frac{1}{a}}t}$ (λ)

Poission Process

$(X_k)_{k\ge1}$ iid ~ exp(λ) — interarrival time

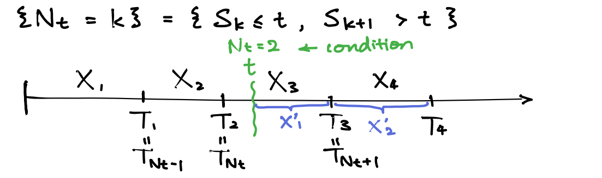


T_k := time of k-th arrival

N_t := num arrivals before time t

~ poi(λt)

$\{N_t = k\} = \{S_k \le t, S_{k+1} > t\}$



Markov prop: ∀t, process after t indep of past

∀n, w.r.t. $\mathbb{P}[\cdot|N_t = n]$, $(X_k')_{k\ge1}$ iid ~ exp(λ)

Equivalently: $\underbrace{N_{t_1}}_{\mathrm{poi}(\lambda t_1)}, \underbrace{(N_{t_2} - N_{t_1})}_{\mathrm{poi}(\lambda(t_2-t_1))}, \dots$ indep …

Fact: Given $N_t = n$, $T_1,\dots,T_n$ unif rand in [0,t]

(pick n rand points then enumerate)

Given point process def, can show (N_t) is Poi(λ) process

- $\mathbb{P}[\Delta N = 1] = \lambda x + o(x)$
- $\mathbb{P}[\Delta N \ge 2] = o(x)$
- $\mathbb{P}[\Delta N = 0] = 1 - \lambda x + o(x)$

$p_k(t+x) = \mathbb{P}[N_t = k, \Delta N = 0] + \mathbb{P}[N_t = k-1, \Delta N = 1]$
$+ \mathbb{P}[N_t = k, \Delta N = 2]$ o(x)

↓ … condition on N_t

$\Rightarrow p_k'(t) = -\lambda p_k(t) + \lambda p_{k-1}(t)$ for k ≥ 1
$-\lambda p_k(t)$ for k = 0

$$\frac{d}{dt}\vec{p}(t) = \begin{bmatrix}p_0'(t)\\p_1'(t)\\p_2'(t)\\p_3'(t)\\\vdots\end{bmatrix} = \begin{bmatrix}-\lambda & 0 & 0 & 0 & \\ \lambda & -\lambda & 0 & 0 & \\ 0 & \lambda & -\lambda & 0 & \cdots\\ 0 & 0 & \lambda & -\lambda & \\ & \vdots & & & \end{bmatrix}\begin{bmatrix}p_0(t)\\p_1(t)\\p_2(t)\\p_3(t)\\\vdots\end{bmatrix}, \quad \begin{cases}p_0(0) = 1\\ p_0(-) = 0\end{cases}$$

Solve ⇒ ∀k ≥ 0, $p_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$

Small-o

f(x) in o(x) as x→0 ⇔ $\frac{f(x)}{x}\xrightarrow[x\to0]{}0$

Differentiability: say $f(t+x) = f(t) + ax - r(x)$

$r(x) := f(t+x) - f(t) - ax$

f differentiable at t with $\frac{df}{dt} = a$ ⇔ r(x) in o(x)

⇔ $f(t+x) = f(t) + ax + o(x)$

$\frac{r(x)}{x} = \left\ulcorner\frac{f(t+x)-f(t)}{x}\right\lrcorner - a$ $\frac{df}{dt}$ at t

LHS → 0 ⇔ RHS → 0

$e^{-\lambda x} = 1 - \lambda x + o(x)$

$\lambda xe^{-\lambda x} = \lambda x + o(x)$

$\mathbb{P}[N_{t+dt} - N_t] = \lambda\,dt$

$\mathrm{Binom}(tn, \frac{\lambda}{n}) \sim \mathrm{Poi}(\lambda t)$ discrete approx of poission process

Markov Property

Poi, Queue, BM all have this

(X_t) has markov prop if (eq defs)
1. Under condition of present $(X_t = \cdot)$, past ⊥ future ← fixed t
2. Future only depend on present out of {past, present}

Strong markov prop is if above still true for rand time (random T_k)

Thm this holds if T_k doesn't depend on own future

call T_k "stopping time" if this is the case

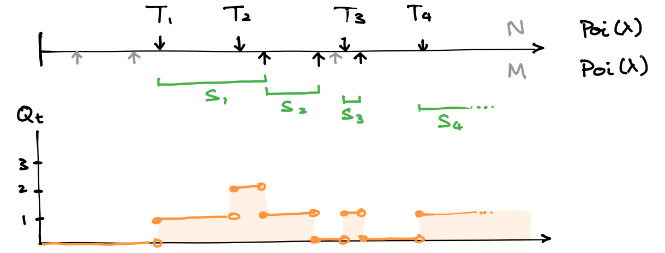
- $L_t(w)$ last visit time to 0 in BM before t
→ No since it depends on future not hitting 0
- $T_a(w)$ first hitting time to a
→ Yes since future doesn't matter
- U_k start service time for k-th customer
→ Yes since once service started we don't care about future

M/M/1 Queue

M - exp enqueue $Q_t(w)$ queue size at time t

M - exp dequeue $T_1,\dots$ Poi(λ) enqueue time μ > λ

1 - one processor $S_1,\dots$ iid exp(μ) service time



Q_t dist $p_k(t) = \mathbb{P}[Q_t = k]$

$p_k(t+x) = \mathbb{P}[\Delta N = 0, \Delta M = 0, Q_t = k]$
$+ \mathbb{P}[\Delta N = 1, \Delta M = 0, Q_t = k-1]$
$+ \mathbb{P}[\Delta N = 0, \Delta M = 1, Q_t = k+1]$
$+ \mathbb{P}[\Delta N = 1, \Delta M = 1, Q_t = k]$
$+ o(x)$ ← ΔM ≥ 2, ΔN ≥ 2 case

$$\frac{d}{dt}\vec{p}(t) = \begin{bmatrix}p_0'(t)\\p_1'(t)\\p_2'(t)\\p_3'(t)\\\vdots\end{bmatrix} = \begin{bmatrix}-\lambda & \mu & 0 & 0 & 0 & \\ \lambda & -(\lambda+\mu) & \mu & \mu & 0 & \\ 0 & \lambda & -(\lambda+\mu) & \mu & 0 & \cdots\\ 0 & 0 & \lambda & -(\lambda+\mu) & \mu & \\ & & \vdots & & & \end{bmatrix}\begin{bmatrix}p_0(t)\\p_1(t)\\p_2(t)\\p_3(t)\\\vdots\end{bmatrix}$$

Suppose $Q_0 \sim \pi$. Then $\vec{p}(0) = \pi$

Solve ⇒ ∃! solution $\vec{p}(t) = [e^{tA}]\,\vec{p}(0)$

∃! π^*, $\vec{p}(t) \to \pi^*$

Set $\vec{p}(0) = \pi^*$ and solve : $\mathbb{P}[\pi^* = k] = (1-\frac{\lambda}{\mu})(\frac{\lambda}{\mu})^k$

So $\pi^* \sim (T-1)$ where $T\sim\mathrm{Geom}(1-\frac{\lambda}{\mu})$

RV Convergence

⓪ $X_n \to X$ surely ⇔ ∀w, $X_n(w) \to X(w)$

① $X_n \to X$ in probability ⇔ ∀δ > 0, $\mathbb{P}[|X_n - X| > 0] \to 0$

② $X_n \to X$ in L^p ⇔ $\mathbb{E}[|X_n - X|^p] = 0$ for fixed p ≥ 0

③ $X_n \to X$ $\mathbb{P}$-almost surely ⇔ $\mathbb{P}[\{w \mid X_n(w) \to X(w)\}] = 1$ ← allows non-empty non-convergence

$$\begin{tikzcd} & ① \arrow[dl, "\not\Leftarrow"'] \arrow[dr, "\not\Rightarrow"] & \\ ② \arrow[rr, "\not\Leftrightarrow"] & & ③ \end{tikzcd}$$

LLN

$(X_n)_{n\ge1}$, $\mathbb{E}X_k \overset{\forall k}{=} m\in\mathbb{R}$, $\mathrm{var}X_k \overset{\forall k}{=} \sigma^2\in\mathbb{R}$, $\mathrm{cov}(X_i,X_j) = 0$

$\Rightarrow \frac{1}{n}\sum_k^n X_k \to m$ in L^2

Also if $(X_n)_{n\ge k}$, $\mathbb{E}X_k \overset{\forall k}{=} m\in\mathbb{R}$, X_i pairwise indep

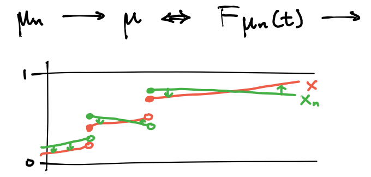
$\Rightarrow \frac{1}{n}\sum_k^n X_k \to m$ in L^1

Dist Convergence

$X_n \xrightarrow[n\to\infty]{d} X \Leftrightarrow F_{X_n}(t) \to F_X(t)$

∀t where $F_X(t)$ doesn't jump

$\mu_n \to \mu \Leftrightarrow F_{\mu_n}(t) \to F_\mu(t)$ ∀t with $\mu(\{t\}) = 0$



If $X_n, X\in\mathbb{Z}$,

$X_n \xrightarrow[n\to\infty]{d} X_n \Leftrightarrow \forall k\in\mathbb{Z}$, $\mathbb{P}[X_n = k] \xrightarrow[n\to\infty]{} \mathbb{P}[X=k]$

These equivalent:

1. $X_n \xrightarrow[n\to\infty]{d} X$

1'. $\mu_{X_n} \xrightarrow[n\to\infty]{d} \mu_X$

2. ∀φ, φ cont & bounded, $\mathbb{E}\varphi(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}\varphi(X)$ ← finite range, bounded top & bottom

2'. ∀t∈ℝ, $\mathbb{E}[e^{itX_n}] \xrightarrow[n\to\infty]{} \mathbb{E}[e^{itX}]$

just with $\varphi(x) = e^{itx} = \cos tx + i\sin tx$, fourier transform of X