2.1-2.2 Sets & Probs

- If S has n elementaries $\Rightarrow$ # event $= 2^n$
- Set properties: ① Communitive / ② Associative / ③ Distributive
  ④ $(A\cup B)^c = A^c \cap B^c$, $(A\cap B)^c = A^c \cup B^c$
- Axioms of Probability:
  ① $P(A) \geq 0$ ② $P(S) = 1$
  ③ If $A\cap B = \emptyset$, then $P(A\cup B) = P(A) + P(B)$
- Extensions from Axiom:
  ① $P(A^c) = 1 - P(A)$ ② $0 \le P(A) \le 1$ ③ $P(\emptyset) = 0$
  ④ $P(A\cup B) = P(A) + P(B) - P(A\cap B)$
  ⑤ three events: $P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)$
  ⑥ $P(A\cup B) \le P(A) + P(B)$ ⑦ If $A \subseteq B$ then $P(A) \le P(B)$

2.3 Counting Methods

- Sampling WITH Replacement, WITH Odering
  # N elements total, select k $= n^k$
- NO Replacement, WITH ordering (P)
  # N total, select out k $= \frac{N!}{(N-k)!}$
  ie. $N\times(N-1)\times(N-2)\times\cdots\times(N-k+1)$
  $\longrightarrow$ Full Permutation: $N!$
- No Replacement, NO Ordering (C)
  # N total, select k $= \binom{N}{k} = \frac{N!}{k!(N-k)!}$
- When split (n) into multiple (k) groups:
  Ex. n balls in bucket, split into categories: n_1 belong to A^{class}, #$B = n_2$, #$C = n_3$ ··· #$K = n_K$ and $n_1 + n_2 + n_3 + \cdots + n_K = n$
  # different ways to PUT BALLS IN A ROW
  $= \binom{n}{n_1 n_2 \cdots n_k} = \frac{n!}{n_1!\, n_2! \cdots n_k!}$ divided by 1 → 2红 3蓝 出现 RBRBB 概率

2.4 Conditional

- Conditional Def: $P(A|B) = \frac{P(A\cap B)}{P(B)}$ for $P(B) > 0$
- Th. Total Probability: $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$
- Baye's Rule:
  $P(B_i|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(A|B_i)\cdot P(B_i)}{\sum_{k=1}^{n} P(A|B_k)\cdot P(B_k)}$

2.5 Independent

- Independent Def: $P(A\cap B) = P(A)P(B)$
  means: $P(A|B) = P(A)$, $P(B|A) = P(B)$
- Pairwise Independent: if A,B,C are Pairwise IND:
  $P(A\cap B) = P(A)P(B)$ ... $A \overset{ind}{\Leftrightarrow} B$, $B \overset{ind}{\Leftrightarrow} C$, $A \overset{ind}{\Leftrightarrow} C$
- Jointly Independent:
  based on previous example, also need:
  $P(A\cap B\cap C) = P(A)P(B)P(C)$

2.6 Sequential Experiments

- Bernoulli: single trial, success or fail
  $P(\text{success}) = p$
- Binomial: n trials, within them, k success
  $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
- Multinomial: n trials, every trial has M classes, result B_j with probability p_j
  probability that every class happen time $(k_1\, k_2 \cdots k_M)$:
  $P(k_1 \cdots k_M) = \frac{n!}{k_1!\, k_2! \cdots k_M!} p_1^{k_1} p_2^{k_2} \cdots p_M^{k_M}$
  → Ex: 扔飞镖 9次，落在区域 1,2,3 概率为 0.2, 0.3, 0.5
  每个区域正好落3次的概率: $\frac{9!}{3!3!3!}(0.2)^3(0.3)^3(0.5)^3$
- Geometric: repeat trials until first success
  P(first success when m^{th} trial happens):
  $P(m) = (1-p)^{m-1}p$, $m = 1, 2, \cdots$
  P(超过k次才成功) $= P(m>k) = (1-p)^k$
- Dependent Experiments
  Ex. two-urn experiment
  $P(s_0, s_1, \ldots, s_n) = P(s_0)P(s_1|s_0)P(s_2|s_1)\cdots P(s_n|s_{n-1})$
- $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$, $|q|<1$
- $\sum_{k=1}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$

3.1-3.2 RandomVariable, Discrete, and pmf

$X: S \to \mathbb{R}$, $z \mapsto X(z)$

- Probability Mass Function (PMF): $p_X(x) = P(X = x)$
  ①②③ Used for pmf valid check
- PMF Props: ① $p_X(x) \ge 0$ ② $\sum_{x\in S_X} p_X(x) = 1$ ③ $P(X\in B) = \sum_{x\in S_B} p_X(x)$
- Valid RV: 对所有 $B\subseteq\mathbb{R}$，事件 $\{z: X(z)\in B\}$ 都属于 $\mathcal{F}$。
  小X在原本S的事件中存在
  $p_X(x) = 0$ for all $x \notin S_X$

3.3&3.5 Expectation & Variance, and Important RVs

- Expected Value:
  $E[X] = \mu_X = \sum_{x\in S_X} x\, p_X(x)$
- E Convergence
  pmf只对**规定好的、有限或可数的一组 x 值**有意义，在这组之外它就当作不存在（或者算 0）。
  $E[|X|] = \sum_k |x_k|\, p_X(x_k) < \infty$
- Expected RV: if $Z = g(X)$ then we have:
  $E[Z] = E[g(X)] = \sum_k g(x_k)\, p_X(x_k)$
- E Formulas:
  if $Z = a g(X) + b h(X) + c$, then: $E[Z] = aE[g(X)] + bE[h(X)] + C$
  → $E[c] = C$
  → $E[aX] = aE[X]$
  → $E[X+c] = E[X] + C$
  → $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$
- Variance: $VAR[X] = E[(X-\mu_X)^2] = \sum_x (x-\mu_X)^2 p_X(x)$
- standard dev: $\sigma_X = \sqrt{VAR(X)}$; $VAR[X] = E[X^2] - (E[X])^2$
- VAR props: $VAR[X+c] = VAR[X]$; $VAR[cX] = c^2 VAR[X]$

分布	pmf	E[X]	Var[X]
Bernoulli(p)	$p^x(1-p)^{1-x}$, x=0,1	p	p(1-p)
Binomial(n,p)	$\binom{n}{k}p^k(1-p)^{n-k}$	np	np(1-p)
Geometric(p)	(1-p)^{k-1}p	1/p	(1-p)/p²
Poisson(λ)	e^{-λ}λ^k/k!	λ	λ
Uniform(1...L)	1/L	(L+1)/2	(L²−1)/12
Zipf(L)	(1/c_L)(1/k)	L/c_L	L(L+1)/(2c_L)−(L/c_L)²

- Poisson
  $P(X=k) = e^{-\lambda}\cdot\frac{\lambda^k}{k!}$
  λ: 平均次数
  k: 实际次数

3.4 Conditional RVs

- Def: $p_{X|C}(x) = P(X = x \mid C) = \frac{P(\{X=x\}\cap C)}{P(C)}$

若 C 本身是"关于 X 的事件"，例如 $C = \{a \le X \le b\}$，则条件 pmf 可直接由 X 的 pmf 算出：

$$p_{X|C}(x) = \begin{cases} \frac{p_X(x)}{P(C)}, & x\in C \\ 0, & x\notin C \end{cases}$$

- Conditional Expectation: $E[X|B] = \sum_x x\, p_{X|B}(x)$
- Conditional VAR: $VAR[X|B] = E[(X - E[X|B])^2|B] = \sum_x (x - E[X|B])^2 p_{X|B}(x)$
- Law of Total Expectation: $E[X] = \sum_i E[X|B_i]\,P(B_i)$ $E[g(X)] = \sum_i E[g(X)|B_i]\,P(B_i)$
- $\sum_{j=0}^{n-1} r^j = \frac{1-r^n}{1-r}$, $|r|<1$

4.1 CDF: $F_X(x) = P(X \le x)$

- Props: $[0 \le F_X(x) \le 1]$, $[\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1]$
  $[$if $a<b$ then $F_X(a) \le F_X(b)]$ $[F_X(b) = \lim_{x\to b^+} F_X(x)]$ right cont'
  $P(a < X \le b) = F_X(b) - F_X(a)$
- to pmf: $p_X(x_k) = P(X = x_k) = F_X(x_k) - F_X(x_k^-)$

4.2 PDF

- $f(x) = \frac{d}{dx}F_X(x)$
- $P(a<x<b) = \int_a^b f_X(x)\,dx$
- $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$
- Props: $[f_X(x) \ge 0]$ $[\int_{-\infty}^{\infty} f_X(x)\,dx = 1]$

4.3 Expected Value of X

- $E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$ Exists only when $\int_{-\infty}^{\infty} |x| f(x)\,dx < \infty$
- $E[X] = \sum_k x_k P(X = x_k)$ for discrete (D)
- if $f_X(m+x) = f_X(m-x)$ then $E[X] = m$
- If $X \ge 0$:
  $E[X] = \int_0^{\infty} (1 - F_X(t))\,dt$; $= \sum_{0}^{\infty} P(X>k)$ (D)
- props: $E\left[\sum_{k=1}^{n} g_k(X)\right] = \sum_{k=1}^{n} E[g_k(X)]$
- Moments: $E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\,dx$

4.4 Imp Continuous RV

- Q-function $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-t^2/2}\,dt = 1 - \Phi(x)$
- $Q(-x) = 1 - Q(x)$
- Normal CDF: $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt$
- $G \Leftrightarrow$ Normal transform: $Z = \frac{X - m}{\sigma}$

表示把任意高斯随机变量 $X \sim \mathcal{N}(m, \sigma^2)$ 标准化，变成标准正态随

- Exp Memoriless: $P(X > t + h \mid X > t) = P(X > h)$

5.8 example:

$$Z = X + Y$$

$$F_Z(z) = P(Z \le z) = P(X + Y \le z) = \iint_{x+y\le z} f_{X,Y}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X,Y}(x,y)\,dy\,dx$$

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \int_{-\infty}^{\infty}\frac{d}{dz}\left(\int_{-\infty}^{z-x} f_{X,Y}(x,y)\,dy\right)dx = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\,dx$$

$$X \perp Y \Rightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y) \Rightarrow \boxed{f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\,dx} \quad \text{(convolution)}$$

假设明天考试有这样一道题：

题目：某个元件的寿命 X 是随机的，均值 $\mu = 100$ 小时，标准差 $\sigma = 10$ 小时。如果我们安装 50 个这样的元件（连续使用，一个坏了换下一个），求总共能撑过 5100 小时的概率。

1. **识别变量**：
   - X_i：第 i 个元件的寿命。
   - $n = 50$。
   - 我们关心的是总寿命 $S_{50} = X_1 + \cdots + X_{50}$。
2. **计算 S_n 的参数**：
   - **总均值**：$E[S_{50}] = 50 \times 100 = 5000$ 小时。
   - **总方差**：$VAR[S_{50}] = 50 \times 10^2 = 5000$。
   - **总标准差**：$STD[S_{50}] = \sqrt{5000} \approx 70.71$。
3. **列出概率表达式**：题目问的是 $P(S_{50} > 5100)$。
4. **标准化 (转化为 Z)**：$Z = \frac{5100 - \text{总均值}}{\text{总标准差}} = \frac{5100-5000}{70.71} = \frac{100}{70.71} \approx 1.41$
5. **得出结论**：$P(Z > 1.41) = Q(1.41)$。(这时候查表即可)

分布	支持集	pdf $f_X(x)$	cdf $F_X(x)$	$E[X]$	$VAR(X)$
Uniform $U[a,b]$	$a \le x \le b$	$\frac{1}{b-a}$	$\begin{cases} 0, & x<a \\ \frac{x-a}{b-a}, & a\le x\le b \\ 1, & x>b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential $\text{Exp}(\lambda)$	$x \ge 0$	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gaussian / Normal $\mathcal{N}(m, \sigma^2)$	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$	$\Phi\left(\frac{x-m}{\sigma}\right)$	m	σ^2
Standard Normal $\mathcal{N}(0,1)$	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	$\Phi(x)$	0	1

(4.5) Functions of RV

已知 X 的 pdf / cdf，给定关系 $Y = g(X)$，求 Y 的 pdf / cdf。

- rule: $P(Y \in C) = P(g(X) \in C) = P(X \in B)$
- 首选方法: $\boxed{F_Y(y) = P(Y \le y) = P(g(X) \le y)}$
- 若 $g(x)$ 严格单调，$y = g(x)$:

$$\boxed{f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|} \quad (x = g^{-1}(y))$$

- 线性变换专用公式（必须秒写）

$$Y = aX + b \Rightarrow \boxed{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)}$$

- 附带结论（Gaussian 专用）:

$$X \sim \mathcal{N}(m, \sigma^2) \Rightarrow Y \sim \mathcal{N}(am + b, a^2\sigma^2)$$

- 多解情形的 pdf 总公式（难点核心）

若 $y = g(x)$ 在 $x_1, \dots, x_n$ 处成立:

$$\boxed{f_Y(y) = \sum_k \frac{f_X(x_k)}{\left|\frac{dy}{dx}\Big|_{x=x_k}\right|}}$$

- Ex. $Y = X^2$, need to get: $P(Y \le y)$

if $y < 0$, $P(Y \le 0) = 0$, $\therefore F_Y(y) = 0 = f(y)$

else: $X^2 \le y \iff -\sqrt{y} \le X \le \sqrt{y}$

CDF

$F_Y(y) = P(-\sqrt{y} \le X \le \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}), \quad y \ge 0$

for pdf of $y>0$: $\frac{d}{dy}F_X(\sqrt{y}) = f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}}$; $\frac{d}{dy}F_X(-\sqrt{y}) = f_X(-\sqrt{y}) \cdot \left(-\frac{1}{2\sqrt{y}}\right)$

$\therefore f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}}, \quad y > 0$

(4.6) Markov & Chebyshev

- Markov: $\boxed{P(X \ge a) \le \frac{E[X]}{a}}$
- Chebyshev: $\boxed{P(|X - M| \ge a) \le \frac{\sigma^2}{a^2}}$

(4.7) Transformation Methods

- Characteristic function:

$$\boxed{\phi_X(v) = E[e^{jvX}] = \int_{-\infty}^{\infty} f_X(x)\, e^{jvx}\, dx}$$

discrete: $\phi_X(v) = \sum_k p_X(k) e^{jvk}$

- Reverse: $p_X(k) = \frac{1}{2\pi}\int_0^{2\pi} \phi_X(v) e^{-jvk} dv$
- Moment: $\boxed{E[X^n] = \frac{1}{j^n}\frac{d^n}{dv^n}\phi_X(v)\Big|_{v=0}}$

(5.1) Two RVs

Example 5.1：抽学生名字 → (身高,体重)

- X(z)=学生身高
- Y(z)=学生体重
- 如事件 {H ≤ 183, W ≤ 82} 对应一个二维区域。

- Example format: $B = \{X \in A_1,\ Y \in A_2\}$

(5.2) Pairs of discrete RVs

- Joint pmf:

$$\boxed{p_{X,Y}(x, y) = P(X = x,\ Y = y)}$$

- $P(B) = \sum_{(x,y)\in B} p_{X,Y}(x, y)$
- Marginal pmf:

$$\boxed{p_X(x) = \sum_y p_{X,Y}(x, y)} \quad \boxed{p_Y(y) = \sum_x p_{X,Y}(x, y)}$$

(5.3) Joint CDF of two RVs

- Joint CDF:

$$\boxed{F_{X,Y}(x, y) = P(X \le x,\ Y \le y)}$$

- Props:

1) If $x_1 \le x_2$, $y_1 \le y_2$ then $F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2)$

2) $\boxed{F_{X,Y}(x, -\infty) = 0, \quad F_{X,Y}(-\infty, y) = 0}$ $\quad \boxed{F_{X,Y}(\infty, \infty) = 1}$

3) $\boxed{F_X(x) = F_{X,Y}(x, \infty) \qquad F_Y(y) = F_{X,Y}(\infty, y)}$

4) $\boxed{\lim_{x \to a^+} F_{X,Y}(x, y) = F_{X,Y}(a, y) \qquad \lim_{y \to b^+} F_{X,Y}(x, y) = F_{X,Y}(x, b)}$

5) $\boxed{P(x_1 < X \le x_2,\ y_1 < Y \le y_2) = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)}$

(5.4) Joint PDF of two Cont RVs

如果 (X,Y) 是“联合连续”（jointly continuous），则它们所有事件的概率都可以通过一个二元 pdf f_{X,Y}(x,y) 来积分求得。

- $\boxed{P[(X, Y) \in B] = \iint_B f_{X,Y}(x, y)\, dx\, dy}$
- $\boxed{F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du}$

  $\boxed{f_{X,Y}(x, y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x, y)}$

- $\boxed{f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx}$

4 协方差的直觉含义

- $COV(X, Y) > 0$: 大的 X 往往伴随大的 Y
- $COV(X, Y) < 0$: 大的 X 往往伴随小的 Y
- $COV(X, Y) \approx 0$: 线性相关性弱

(5.5) Independent RVs

两个随机变量 X 与 Y 独立，当且仅当：
任意关于 X 的事件与任意关于 Y 的事件都独立。
其等价条件是：
联合分布 = 边缘分布的乘积。

- $\boxed{f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)}$

  $\boxed{P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)}$

- $\boxed{F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)}$

(5.6) Joint Moments & E of 2 RVs

- Expected Value:

$$\boxed{E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)\, f_{X,Y}(x, y)\, dx\, dy}$$

$E[g(X, Y)] = \sum_i\sum_j g(x_i, y_j)\, p_{X,Y}(x_i, y_j)$ discrete

- If X, Y Independent: $(g(X, Y) = g_1(X)\, g_2(Y))$

$$\boxed{E[g_1(X) g_2(Y)] = E[g_1(X)]\, E[g_2(Y)]}$$

- Joint Moments: the (j,k)-th: $\boxed{E[X^j Y^k]}$

orthogonal if $E[XY] = 0$

- Covariance:

$$\boxed{COV(X,Y) = E[(X - E[X])(Y - E[Y])]}$$

$$\boxed{COV(X,Y) = E[XY] - E[X]E[Y]}$$

- If X, Y Independent:

$E[XY] = E[X]E[Y] \Rightarrow \boxed{COV(X, Y) = 0}$

- ↳ CANNOT REVERT!
- Correlation Coefficient: $\boxed{\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}}$

$\sigma_X = \sqrt{VAR(X)}, \quad \sigma_Y = \sqrt{VAR(Y)}$ $\qquad -1 \le \rho_{X,Y} \le 1$

(5.7) Cond' prob & Conditional E

- 离散情型: $\boxed{p_{Y|X}(y \mid x) = \frac{p_{X,Y}(x, y)}{p_X(x)}}$

$p_{X,Y}(x, y) = p_{Y|X}(y \mid x)\, p_X(x)$

$P(Y \in A) = \sum_x P(Y \in A \mid X = x)\, p_X(x)$

- If X, Y independent: $\boxed{p_{Y|X}(y \mid x) = p_Y(y)}$
- 连续情型: conditional pdf: $\boxed{f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)}}$

conditional cdf: $\boxed{F_{Y|X}(y \mid x) = \int_{-\infty}^{y} f_{Y|X}(t \mid x)\, dt}$

total prob's continuous version: $\boxed{P(Y \in A) = \int_{-\infty}^{\infty} P(Y \in A \mid X = x)\, f_X(x)\, dx}$

- cond' cont' E: $\boxed{E[Y \mid X = x] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y \mid x)\, dy}$ $\quad \Sigma$ if discont'

cond' E is a function: $\boxed{E[Y \mid X = x] = g(x)}$ 于是我们可以定义随机变量: $\boxed{E[Y \mid X]}$

- Law of total E: $\boxed{E[Y] = E[E[Y \mid X]]}$

if 例 $Y = aX + b + N, \quad N \perp X,\ E[N] = 0$ then $\boxed{E[Y \mid X] = aX + b}$

- 若 X, Y 独立: $\boxed{E[Y \mid X] = E[Y]}$

(5.8) Functions of 2 RVs

5.8 研究 由 (X, Y) 生成的新随机变量（一个或两个）如何求分布。

1. 一个函数 $Z = g(X, Y)$
2. 两个函数 $(Z_1, Z_2) = (g_1(X, Y), g_2(X, Y))$
3. 线性变换 的 pdf（Jacobian）

1 万能思路（最稳）

$$\boxed{F_Z(z) = P(Z \le z) = P(g(X, Y) \le z)}$$

把不等式变成 (x, y) 平面上的区域 R_z，对联合 pdf 积分：

$$F_Z(z) = \iint_{R_z} f_{X,Y}(x, y)\, dx\, dy$$

再对 z 求导得 $f_Z(z)$。

三、条件法（降维神器）

若已知 $Y = y$，则 $Z = g(X, y)$ 是一维问题，可用 4.5 的方法先求

$$f_{Z|Y}(z|y)$$

再“平均回去”：

$$\boxed{f_Z(z) = \int f_{Z|Y}(z|y)\, f_Y(y)\, dy}$$

直觉：先在每个切片里算，再对切片加权平均。

2. 两个函数

1 CDF 法（通用）

$$\boxed{F_{Z_1,Z_2}(z_1, z_2) = P(g_1(X, Y) \le z_1,\ g_2(X, Y) \le z_2)}$$

把条件翻译成 (x, y) 平面区域，对 $f_{X,Y}$ 积分。

$\begin{cases} V = aX + bY \\ W = cX + eY \end{cases}$ $\quad A^{-1} = \frac{1}{\det A}\begin{pmatrix} e & -b \\ -c & a \end{pmatrix}$ $\quad \boxed{ae - bc}$

- Jacobian: set: $\begin{pmatrix} V \\ W \end{pmatrix} = A\begin{pmatrix} X \\ Y \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ c & e \end{pmatrix}$, $\det A \ne 0$

$$\boxed{f_{V,W}(v, w) = \underbrace{f_{X,Y}(A^{-1}(v, w))}_{f_{X,Y}(x,y)} \cdot \frac{1}{|\det A|}}$$

假设你有两个旧变量 (X, Y)，通过变换变成了两个新变量 (U, V)：
$U = g_1(X, Y) \quad V = g_2(X, Y)$

$$\mathbf{J} = \begin{bmatrix} \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y} \\ \frac{\partial V}{\partial X} & \frac{\partial V}{\partial Y} \end{bmatrix}$$

[公式右边的 x 和 y 必须**全部反解**并替换成 u 和 v。不能留着旧变量在结果里！] important!

(7.1) Sum of RVs

$S_n = X_1 + X_2 + \cdots + X_n \quad E[S_n] = E[X_1] + \cdots + E[X_n]$

- VAR — two RVs: $Z = X+Y$; $VAR(Z) = E[(Z - E_Z)^2]$ $\boxed{\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)}$
- VAR — n variables: $\mathrm{Var}(S_n) = \sum_{k=1}^{n}\mathrm{Var}(X_k) + \sum_{j \ne k}\mathrm{Cov}(X_j, X_k)$ if 相互独立: $\boxed{\mathrm{Var}(S_n) = \sum_{k=1}^{n}\mathrm{Var}(X_k)}$
- VAR - if $X_k \overset{\text{iid}}{\sim} (\mu, \sigma^2)$ $\quad \boxed{E[S_n] = n\mu,} \quad \boxed{\mathrm{Var}(S_n) = n\sigma^2}$
- 已知 $X_1, \dots, X_n$ 独立，如何求 S_n 的 pdf?

1 两个变量：卷积（你在 5.8 已经见过）

若 $Z = X + Y$，且 X, Y 独立: $\boxed{f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z - x)\, dx}$

2 多个变量：特征函数是王者 $\mathcal{L}_X(v) = E[e^{jvX}]$

$\boxed{\mathcal{L}_{X+Y}(v) = \mathcal{L}_X(v)\, \mathcal{L}_Y(v)}$ 对 S_n: $\mathcal{L}_{S_n}(v) = \prod_{k=1}^{n}\mathcal{L}_{X_k}(v)$

(7.2) Sample Mean & Law of Large Num

$X_1, X_2, \dots, X_n$ iid, $\quad E[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2$

- 样本均值 Sample Mean $\quad M_n = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \frac{1}{n}S_n$

$M_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$

- 均值与方差: $\boxed{E[M_n] = \mu} \quad \boxed{\mathrm{Var}(M_n) = \frac{\sigma^2}{n}}$
- chebichev 在这里: $\boxed{P(|M_n - \mu| \ge \varepsilon) \le \frac{\sigma^2}{n\varepsilon^2}}$
- WLLN: $\boxed{\lim_{n\to\infty} P(|M_n - \mu| \le \varepsilon) = 1}$ $\quad \boxed{P(|M_n - \mu| > \varepsilon) \xrightarrow[n\to\infty]{} 0}$
- SLLN: $\boxed{P\left(\lim_{n\to\infty} M_n = \mu\right) = 1}$

(7.3) Central Limit Theorem

- $\boxed{Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(M_n - \mu)}{\sigma}}$ $\quad E[Z_n] = 0 \quad \mathrm{Var}(Z_n) = 1$
- CLT: $\boxed{Z_n \xrightarrow{d} \mathcal{N}(0, 1)}$ $\quad$ equal writing: $\boxed{S_n \approx \mathcal{N}(n\mu,\ n\sigma^2)}$ $\quad \boxed{M_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)}$
- 概率近似模板: $\boxed{P(S_n \le a) \approx \Phi\left(\frac{a - n\mu}{\sigma\sqrt{n}}\right)}$

$\boxed{P(M_n \le m) \approx \Phi\left(\frac{\sqrt{n}(m - \mu)}{\sigma}\right)}$

Z_n 是:
“样本均值偏离真实均值多少个‘标准差单位

X Y 都是正态, their joint pdf:

$$f_{X,Y}(x, y) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}}e^{-y^2/2} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$$

$$\mu_{X|Y} = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y) \qquad \sigma^2_{X|Y} = \sigma^2_X(1 - \rho^2)$$

$$P(W > x) = Q\left(\frac{x - \mu}{\sigma}\right)$$

$$\frac{d}{dx}\sin(x^2) = \cos(x^2) \cdot 2x$$

- MSE: $E[(\hat{X} - X)^2]$