Relative Freq.

$f_k = \frac{n_k}{n}$

1) $f_k \geq 0$
2) $\sum_k f_k = 1$

$P(A) = \frac{\#\ \xi \text{ in } A}{\text{total } \#\ \xi}$

Permutation: arrange n objects at k positions (order matters)

without replacement: $\frac{n!}{(n-k)!}$ with replacement: n^k

Combination: choose k objects from n objects (order doesn't matter)

without replacement: $\frac{n!}{k!(n-k)!} = \binom{n}{k}$ "n choose k" with replacement: $\frac{(n+k-1)!}{k!(n-k)!}$

$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, |x|<1$

$\sum_{k=a}^{\infty} x^k = \frac{x^a}{1-x}$

$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$

Ax. of Prob.

1) $P(A) \geq 0$
2) $P(S) = 1$
3) if $A \cap B \neq 0$, $P(A \cup B) = P(A) + P(B)$

Cond. Prob.

$P(A|B) = \frac{P(A\cap B)}{P(B)}$

$P(A\cap B) = P(A|B)P(B) = P(B|A)P(A)$

Baye's Rule

$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$

Total Prob. Law

$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$

Indep. Events Jointly vs. Pairwise

$P(A|B) = P(A)$

$P(A\cap B) = P(A)P(B)$

Ax. of RVs

1) $p_X(x_k) \geq 0$
2) $\sum_k p_X(x_k) = 1$
3) if $B = \{x^{(1)}, x^{(2)}, \dots\} \subset S$ then $P(B) = \sum_k p_X(x^{(k)})$

PMF: Prob. Mass Func^{n} Prob that X is x_k

$p_X(x_k) = P[X = x_k] = P[\{\xi : X(\xi) = x_k\}]$

CDF: Cumul. Distr. Func^{n} Prob that X is $\leq x$

$F_X(x) = P[X \leq x] = P[\{\xi : X(\xi) \leq x\}]$

PDF: Prob. Distr. Func^{n} Prob that X is near x (prob of being in small segment around x)

$f_X(x) = \frac{d}{dx} F_X(x)$

Prop. of Prob.

1) $P(A) \leq 1$
2) $P(A^c) = 1 - P(A)$
3) $P(\emptyset) = 0$
4) if $A_i \cap A_j = \emptyset \ \forall i \neq j$, then $P(\bigcup_i^{\infty} A_i) = \sum_i^{\infty} P(A_i)$
5) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

Expecta^{n} Values:

$E[X] = \sum_k x_k\, p_X(x_k) = m_X$

$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x,y)\,dx\,dy$

Variance:

$\sigma_X^2 = E[D^2] = E[(\underbrace{X - m_X}_{})^2]$ deviation $D = X - m_X$

1) $\sigma_X^2 \geq 0$
2) if $\sigma_X^2 = 0$, then $X = E[X] = \text{constant}$
3) if $Y = X + \text{constant}$, then $\sigma_Y^2 = \sigma_X^2$
4) if $Y = aX$, then $\sigma_Y^2 = a^2\sigma_X^2$
5) $\sigma_X^2 = E[X^2] - (E[X])^2$

Bernouli can also be Indicator RV

$P(\text{success}) = p = P[A]$

$P(\text{failure}) = 1 - p$

$E[X] = p$

$\sigma_X^2 = p(1-p)$

Binomial

$P\left(\substack{k \text{ success} \\ \text{in } n \text{ trials}}\right) = \binom{n}{k} p^k (1-p)^{n-k}$

$E[X] = np$

$\sigma_X^2 = np(1-p)$

Geometric

$P\left(\substack{k \text{ trials for} \\ 1^{st} \text{ success}}\right) = (1-p)^{k-1} p$

$E[X] = \frac{1}{p}$ $\sigma_X^2 = \frac{1-p}{p^2}$

Poisson approximates Binomial n large, p small

Uniform RV

$S_X = \{1, 2, 3, \dots, L\}$

$p_X(k) = \frac{1}{L}$

$E[X] = \frac{L+1}{2}$

$\sigma_X^2 = \frac{L^2-1}{12}$

Poisson RV

$\alpha = np$, $\lambda = np$, $\alpha = \lambda T$

P(success)/time unit; prob of event per time unit; time period

$p_X(k) = e^{-\alpha} \cdot \frac{\alpha^k}{k!}$, $k = 0, 1, 2, \dots$

$E[X] = \alpha$ $\sigma_X^2 = \alpha$

CDF Prop.

1) $0 \leq F_X(x) \leq 1$
2) $\lim_{x\to\infty} F_X(x) = 1$
3) $\lim_{x\to-\infty} F_X(x) = 0$
4) CDF is cont. from the right
5) $P[X=x] = F_X(x) - \lim_{\substack{h\to 0 \\ h>0}} F_X(x-h)$ discont. at x (value at x minus value just b4)
6) if $F_X(x)$ is cont, then for any x: $P[X=x] = 0$ no discont.
7) $P[a < X \leq b] = F_X(b) - F_X(a)$
8) $F[X > x] = 1 - F_X(x)$
9) $F_X(x)$ is a non-decreasing function of x

- discrete RV ↔ piecewise constant CDF
- continuous RV ↔ continuous CDF
- mixed RV ↔ piecewise continuous CDF

PDF Prop.

1) $f_X(x) \geq 0$
2) $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$
3) $P[a < X \leq b] = \int_a^b f_X(t)\,dt$
4) $\int_{-\infty}^{\infty} f_X(x) = 1$

Given PDF of X and $Y = g(X)$, PDF of Y:

$f_Y(y) = \sum_{\substack{x_k \text{ s.t.} \\ g(x_k) = y}} \frac{f_X(x_k)}{\left|\left[\frac{dy}{dx}\right]_{x=x_k}\right|}$

$\Phi_X(j\omega) = E[e^{j\omega X}] = \int_{-\infty}^{\infty} f_X(x) e^{j\omega x}\,dx = \sum_k p_X(x_k) e^{j\omega x_k}$

$\frac{d^n}{d\omega^n} \Phi_X(j\omega)\Big|_{\omega=0} = j^n E[X^n]$

Exponential RV

$f_X(x) = \lambda e^{-\lambda x} u(x)$

$E[X] = \frac{1}{\lambda}$ $\sigma_X^2 = \frac{1}{\lambda^2}$

characterizes inter-event times of a Poisson $\alpha = \lambda t$

$P[T > t] = P\left[\begin{matrix}\text{no events}\\ \text{in time } t\end{matrix}\right]$

$= e^{-\alpha}\frac{(\alpha)^0}{0!} = e^{-\lambda t}u(t)$

$F_X(x) = (1 - e^{-\lambda t})u(t)$

Gaussian RV

$f_X(x) = (2\pi\sigma^2)^{-\frac{1}{2}}\exp\left[\frac{-(x-m)^2}{2\sigma^2}\right]$

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

$X \sim \mathcal{N}(m, \sigma^2)$ X is normal with mean m variance σ^2

$\frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$ standard normal

$\Phi(x) = P[X \le x] = \int_{-\infty}^{x} \mathcal{N}(0,1)\,dt$

$Q(x) = P[X \ge x] = \int_{x}^{\infty} \mathcal{N}(0,1)\,dt$

$Q(-x) = 1 - Q(x)$

Uniform RV

$f_X(x) = \frac{1}{\text{length}} = \frac{1}{\text{area}}$

$= \frac{1}{b-a}$, $a < x < b$

$E[X] = \frac{a+b}{2}$ $\sigma^2 = \frac{(b-a)^2}{12}$

Laplacian RV

$f_X(x) = \frac{\alpha}{2}e^{-\alpha|x|}$

$\Phi_X(j\omega) = \frac{\alpha/2}{\alpha - j\omega} + \frac{\alpha/2}{\alpha + j\omega}$

$E[X] = 0$ $\sigma^2 = \frac{2}{\alpha^2}$

Markhov (+ve RVs)

$P[X \ge a] \le \frac{E[X]}{a}$

Works best for large a. If mean is small, then prob of being much greater than mean is small

Chebyshev (any RV)

$P[|X - m_X| \ge a] \le \frac{\sigma_X^2}{a^2}$

If variance is small, then prob that std deviation is large is small

Joint PMF

Joint PMF: $P_{XY}(x_k, y_j) = P[X = x_k, Y = y_j]$

marginal PMFs

$\sum_{(x_k, y_j) \in S_{XY}} P_{XY}(x_k, y_j) = 1$

$P_X(x_k) = \sum_{y_j} P_{XY}(x_k, y_j)$

$P_Y(y_j) = \sum_{x_k} P_{XY}(x_k, y_j)$

$P[B] = \sum_{(x_k, y_j) \in B} P_{XY}(x_k, y_j)$

Correlation Coefficient

$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$

$\sqrt{\mathrm{Var}(X)}$ $\sqrt{\mathrm{Var}(Y)}$

1 strong +ve correlation if $X\uparrow, Y\uparrow$

weakly cpld

0

strongly cpld

strong −ve correlation if $X\uparrow, Y\downarrow$

−1

LMMSE

LMMSE: $Y = X + N$ X, N are zero mean

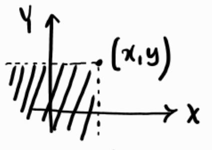
Estimate X with $\hat{X} = aY$

s.t. $\min_a E[(X - \hat{X})^2]$

$a = \frac{\mathrm{cov}(X,Y)}{\sigma_Y^2} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_N^2}$

Joint CDF

Joint CDF: $F_{XY}(x,y) = P[X \le x, Y \le y]$



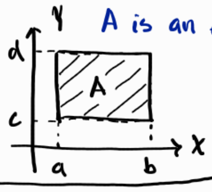
CDF of points above and to the right are always greater

marginal CDFs

$F_X(x) = \lim_{y \to \infty} F_{XY}(x,y)$ discard regions of small y

$F_Y(y) = \lim_{x \to \infty} F_{XY}(x,y)$ discard regions of small x

A is an event of product form



$P[A] = F_{XY}(b,d) - F_{XY}(a,d) - F_{XY}(b,c) + F_{XY}(a,c)$

Covariance

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

covariance of X and Y $= E[XY] - E[X]E[Y]$

If $\mathrm{cov}(X,Y) = 0$, X, Y are uncorrelated

$E[XY]$: correlation of X and Y

If $E[X,Y] = 0$, X, Y are orthogonal

Indep. RVs are uncorrelated : $\mathrm{cov}(X,Y) = 0$ but uncorrelated RVs need not be indep

$E_Y[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy \to E_Y[Y/x] = \int_{-\infty}^{\infty} y f_Y(y/x)\,dx$

$E[X] = \sum_i E[X/B_i]P[B_i]$

$E[g(Y)] = E_X\left[E_Y[g(Y)/x]\right]$

5.7: Conditional Prob

1) Both discrete $P_Y(y_j / X = x_k) = \frac{P_{XY}(x_k, y_j)}{P_X(x_k)}$

$P_Y(y_j) = \sum_{x_k} P_Y(y_j / X = x_k) P_X(x_k)$

2) $Y = X + N$, $X \in \{-A, +A\}$, $N \sim \mathcal{N}(0, \sigma^2)$

$Y/X = -A \sim \mathcal{N}(-A, \sigma^2)$

$Y/X = +A \sim \mathcal{N}(+A, \sigma^2)$

3) Both Cont. : $f_Y(y/X = x) = \frac{f_{XY}(x,y)}{f_X(x)}$

$f_Y(y) = \int_{-\infty}^{\infty} f_Y(y/X = x) f_X(x)\,dx$

Joint PDFs

Joint PDFs : Prob that (X,Y) lands in small region around (x,y)

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

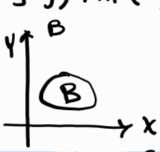
$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y) = 1$

marginal PDFs

$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$

$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$

$P[B] = \iint f_{XY}(x,y)\,dA$



Iff X, Y indep: $f_{XY}(x,y) = f_X(x) f_Y(y)$

Transformations

$\begin{bmatrix} W \\ Z \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}$ given $f_{XY}(x,y)$, find $f_{WZ}(w,z) = \frac{f_{XY}(x,y)}{|ad - bc|}$ $\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1}\begin{bmatrix} w \\ z \end{bmatrix}$

$Z = g(X,Y)$ find $f_Z(z)$

If $Z = X + Y$ and X, Y indep, then $\Phi_Z(\omega) = \Phi_X(\omega)\Phi_Y(\omega)$

Step ①: condition on one. Let $X = x$: $Z = g(x, Y)$. This gives us $f_Z(z / X = x)$

Step ②: Average it (x here) $f_Z(z) = \int_{-\infty}^{\infty} f_Z(z/x) f_X(x)\,dx$

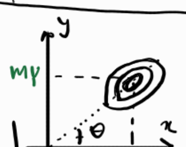
$\begin{pmatrix} W \\ Z \end{pmatrix} = g\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{bmatrix} g_1(X,Y) \\ g_2(X,Y) \end{bmatrix}$ $J = \begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{bmatrix}$

$f_{WZ}(w,z) = \left[\frac{f_{XY}(x,y)}{|\det J|}\right]_{(x,y) = g^{-1}(w,z)}$

$X \sim \mathcal{N}(m_X, \sigma_X^2)$

$Y \sim \mathcal{N}(m_Y, \sigma_Y^2)$

$\theta = \mathrm{atan}\left(\frac{\rho_{XY}\sigma_X\sigma_Y}{\sigma_X^2 - \sigma_Y^2}\right)$



$Y/X = x \sim \mathcal{N}\left(\rho_{XY}\frac{\sigma_Y}{\sigma_X}(x - m_X) + m_Y, \sigma_Y^2(1 - \rho_{XY}^2)\right)$

5.6: Joint Moments

5.6: Joint Moments : $E[X^k Y^j]$ moment, $E[(X - m_X)^k (Y - m_Y)^j]$ central moment

1) $E[X + Y] = E[X] + E[Y]$ always

2) $E[g_1(X) \cdot g_2(Y)] = E[g_1(X)] \cdot E[g_2(Y)]$ only if X, Y are indep.

3) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \underbrace{2E[(X - m_X)(Y - m_Y)]}_{\text{crossterm}}$ always

If X, Y indep, then crossterm = 0

$\mathrm{Var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$ only if X, Y indep.

Sums of RVs

$S_n = \sum_{i=1}^{n} X_i$, $E[S_n] = \sum_{i=1}^{n} E[X]$

$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$ $\forall X_i$ uncorrelated

$\Phi_{S_n} = \prod_{i=1}^{n} \Phi_{X_i}$ $\forall X_i$ indep.

$\Phi_{S_n} = [\Phi_{X_i}]^n$ for iid RVs X_i

Sample mean : $M_n = \frac{1}{n}S_n$

$E[M_n] = E[X]$

$\mathrm{Var}(M_n) = \frac{\mathrm{Var}(X)}{n}$

If S_n is sum of n iid RVs X :

$E[S_n] = nE[X]$ $\mathrm{Var}(S_n) = n\mathrm{Var}(X)$

$E[I(A)] = P[A]$

S_n = # times A happens

M_n = relative freq. of A

CLT: works for $n \ge 30$

$S_n \sim \mathcal{N}(nm_X, n\sigma_X^2)$

$M_n \sim \mathcal{N}(m_X, \frac{\sigma_X^2}{n})$

$P[|M_n - P[A]| < \epsilon] \ge 1 - \frac{\mathrm{var}(I(A))}{n\epsilon^2}$