EE2012 Final cheatsheet by wjb75

1 Chap 1

Probability as a

1. subjective quantitative belief
2. Relative freq demended by law of large numbers $P(A) = lim_{n\to\infty} f_h(n) = lim_{n\to\infty} \frac{N_h}{n}$

1.1 Axioms of Probability

1. $0 \le P(A_i) \le 1 \; \forall A_i$
2. $P(S) = 1$
3. for any series of mutually exclusive events E_i,

$$P(\cup_{i=0}^{\infty} E_i) = \sum_{i=0}^{\infty} P(E_i)$$

note that individual outcomes are by default mutually exclusive events

1.2 Properties of Probability

1. $P(E^c) = 1 - P(E)$
2. $P(F) = P(EF) + P(E^cF)$
3. $E \subseteq F \implies P(E) \le P(F)$
4. $P(E \cup F) = P(E) + P(F) - P(EF)$
5.

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^{n} P(E_i) - \sum_{1\le i_1<i_2\le n} P(E_{i_1} \cap E_{i_2})$$
$$+ \sum_{1\le i_1<i_2<i_3\le n} P(E_{i_1} \cap E_{i_2} \cap E_{i_3}) - \cdots$$
$$+ (-1)^{r+1} \sum_{1\le i_1<i_2<\cdots<i_r\le n} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_r}) + \cdots$$
$$+ (-1)^{n+1} P(E_1 \cap E_2 \cap \cdots \cap E_n).$$

1.3 Combinatorics

Dont forget to use graphical methods

Die 1 \ Die 2	1	2	3	4	5	6
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)

1. n total Permutation $n!$
2. k-permutation from n :$nPk = \frac{n!}{(n-k)!}$
3. combination of n from k $\binom{n}{k} = \frac{n!}{(n-k)!(k!)}$
4. partition/ multinomial coefficient / not all distinct permutations : $\binom{n}{m1, m2, m3, \ldots} = \frac{n!}{m1!m2!m3!\ldots}$
5. pascal identity $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$
6. binomial theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

so $\sum_{k=0}^{n} \binom{n}{k} = 2^n, \sum_{k=0}^{n} \binom{n}{k}(-1)^k = 0$

7. the equation $x_1 + x_2 + \ldots + x_r = n, x_i > 0 \forall i$ has $\binom{n-1}{r-1}$ positive integer solutions
8. the equation $x_1 + x_2 + \ldots + x_r = n, x_i \ge 0 \forall i$ has $\binom{n+r-1}{r-1}$ nonnegative integer solutions
9. vandermonde's identity

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$$

10. hockey stick identity

$$\sum_{k=r}^{n} \binom{k}{r} = \binom{n+1}{r+1}$$

11. derangement

$$!n = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}$$

2 Chap 2 Conditional Probability and Bayes Rule

1. $P(A|B) = \frac{P(AB)}{P(B)}$ if $P(B) \ne 0$
2. $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ if $P(B) \ne 0$
3. law of total probability

$$P(A) = \sum_i P(A|E_i)P(E_i)$$

where E_i forms a partition of the sample space (mutually exclusive and union is S)

4. Two events A and B are independent iff $P(AB) = P(A)P(B)$ In conditionnal probability, if $P(B|A) = P(A)$, $P(A) \ne 0$

2.1 Markov Chain

1. probability to absorption states to reach a specific absorption state s, the probabilities from each state $a_1, a_2, .., a_m$ are unique solution to $a_s = 1, a_i = 0 \forall absorption\ state\ i$ and $a_i = \sum_{j=1}^{M} a_j p_{ij}$
2. expected time to absorption

$$t_i = 1 + \sum_{j=1}^{M} t_j p_{ij}$$

where the 1 is added for each step increment

3 Chap 3 Discrete Random Variable

1. the PMF of a random variabe X is
$$P_X(k) = \begin{cases} P(X=k), & \text{if } k \in S_x, \\ 0, & otherwise \end{cases}$$
where S_x is the support of X
2. properties of PMF
   (a) $0 \le P_X(i) \le 1$
   (b) $\sum_{i=0}^{\infty} P_X(i) = 1$
3. CDF $F_X(x) = P(X \le x)$ for $x \in \mathbb{R}$
4. expectation
$$E[X] = \sum_{\text{all x}} x P_X(x)$$
Expectations are linear, regardless of independencies
5. Moments
$E[X^n] = \sum_{\text{all x}} x^n P_X(x)$ is called the n^{th} moment of X
$$E(X^n) \ge [E(X)]^n$$
(Jensen's inequality)
6. Variance and standard deviation
$$Var[X] = E((X - E[X])^2) = E[X^2] - E[X]^2$$
$$\sigma_X = \sqrt{Var[X]}$$
$$Var[aX + b] = a^2 Var[X]$$
$Var[X+Y] = Var[X] + Var[Y]$ iff X and Y are independent

3.1 Discrete Distribution.

1. Bernoulli
PMF: $P(X = 1) = p, P(X = 0) = 1 - p$
CDF: $F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$
Mean : p, Variance: $p(1-p)$
2. Binomial PMF : $\binom{n}{k} p^k (1-p)^{n-k}$
CDF: $F(k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$
Mean : np, Variance: $np(1-p)$
3. Uniform PMF: $\frac{1}{b-a+1}$
CDF: $F(x) = \frac{x-a+1}{b-a+1}$
Mean: $\frac{a+b}{2}$, Variance: $\frac{(b-a+1)^2-1}{12}$
4. Geometric PMF: $(1-p)^k p$
CDF: $F(k) = 1 - (1-p)^{k+1}$
Mean : $\frac{1}{p}$, Variance: $\frac{1-p}{p^2}$ note that for geometric distribution $P(X > k) = (1-p)^k$
5. Poisson PMF: $\frac{e^{-\lambda}\lambda^k}{k!}$
CDF: $F(k) = e^{-\lambda} \sum_{i=0}^{k} \frac{\lambda^i}{i!}$
Mean: λ, Variance: λ
6. Hypergeometric(N,K,n)
PMF: $\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$
Mean: $\frac{nK}{N}$, Variance: $\frac{nK(N-K)(N-n)}{N^2(N-1)}$

4 Chap 4. Multiple Discrete Random Variable

4.0.1 Joint PMF

(X,Y) with support $S_{X,Y} = \{(x_i, y_i),,.j = 1,2,...\}$
$\forall(x,y) \in S_{X,Y}, A := \{X = x\} \bigcap \{Y = y\}$
$P_{X,Y}(x,y) := P(A) = P(X = x, Y = y)$

4.0.2 Marginal PMF

$P(X = x_i) = \sum_{y_i \in S_Y} P_{X,Y}(x_i, y_i)$

4.0.3 Expected value

$E[g(X,Y)] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x,y) P_{X,Y}(x,y)$

4.0.4 Independence

Necessary and Sufficient Condition
$P_{X,Y}(x,y) = P_X(x) P_Y(y) \quad \forall(x,y) \in S_{X,Y}$
Determinisitc functions of independent RVs are Independent
Independence $\rightarrow\not\leftarrow$Correlation=0

4.1 Covariance and Correlation (same for Continuous

$Cov[X,Y] := E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$
$Cov[X,X] = Var[X]$
$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \in [-1,1]$ where $\sigma_X = \sqrt{Var[X]}$

4.2 Sum of Multiple RV

4.2.1 Sum as Discrete Convolution

$Z = X + Y \ P_Z j = \sum_{k \in S_X,\ j-k \in S-Y} P_X(k) P_Y(j-k)$
one of the distribution is flipped before shifting by j units

4.2.2 Expectation Linearity

$E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i]$ regardless of X_i independence or not

4.2.3 Variance

$Var[X+Y] = Var[X] + Var[Y] + 2Cov(X,Y)$
$Var[\sum_{i=1}^N X_i] = \sum_{i=1}^N Var[X_i] + \sum_{\{(i,j)|i \neq j\}} Cov(X_i, X_j)$
note that i,j are not in increasing sequence and are interchangeable

4.2.4 Reduction of Mean Variance through Independent Sampling

$S_n = \sum_{i=1}^n X_i$
$Var[S_n] \propto n$
$\bar{X} = \frac{Sn}{n}$
$Var[\bar{X}] = Var[\frac{Sn}{n}] = \frac{\sigma^2}{n} < \sigma^2$

5 Continuous Random Variable

5.1 PDF

$P(X \in (a,b)) = \int_a^b f_x(t) dt$

1. $f_x > 0 \forall x \in S$
2. $\int_S f_X(x) dx = 1$ (normalization)
3. $\forall A, P(X \in A) = \int_A f_x(x) dx$
4. $F_X(a) = P(X \leq a) = \int_{-\infty}^a f_X(\tau) d\tau$

5.2 CDF

1. $0 \leq F_X(x) \leq 1, \forall x \in \mathbb{R}$
2. $$\lim_{x \to -\infty} F_X(x) = 0, \lim_{x \to \infty} F_X(x) = 1$$
3. $a < b \Rightarrow F_X(a) \leq F_X(b)$
4. $P(a < X \leq b) = F_X(b) - F_X(a)$

5.3 Expectation

$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$

1. If $f_X(x)$ is even symmetry about m, $E[X] = m$
2. If X is non negative RV,
$$E[X] = \int_0^\infty (1 - F_X(x)) dx = \int_0^\infty P(X > x) dx$$
3. for any function g where $g(x) \geq 0 \forall x$,
$$E[g(X)] = \int_0^\infty (1 - F_g(x)) dx = \int_0^\infty P(g(X) > x) dx$$

5.4 Continuous Distributions

5.4.1 Continuous Uniform $X \sim U(a,b)$

$f_x(x) = \frac{1}{b-a}, a < x < b \ E[X] = \int_a^b x f_X(x) dx = \frac{b+a}{2}$
$E[X^2] = \int_a^b x^2 f_X(x) dx = \frac{b^2+ab+a^2}{3} \ Var[X] = \frac{(b-a)^2}{12}$

5.4.2 Gaussian/Normal RV $X \sim N(\mu, \sigma^2)$

$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2}), x \in \mathbb{R}$
Conversion to standard gaussian $Z = \frac{X-\mu}{\sigma} \sim N(0, 1^2)$
Non-correlation $\Leftrightarrow$ Independence of 2 Gaussian RV

5.4.3 Exponential $T \sim Exp(\lambda)$

$f_T(t) = \lambda e^{-\lambda t}, t > 0 \ F_T(t) = 1 - e^{-\lambda t}, t > 0$
$P(T > t) = 1 - F_T(t) = e^{-\lambda t}, t > 0 \ E[T] = \frac{1}{\lambda}, Var[T] = \frac{1}{\lambda^2}$
$P(T > t_1 + t_2 | T > t_1) = P(T > t_2)$

6 Chapter 6 Multiple Continuous random variable

6.1 Sum as Convolution

$f_{X+Y}(z) = f_X(z) * f_Y(z) = \int_{-\infty}^{\infty} f_X(t) f_Y(z-t) dt$

6.2 Central Limit Theorem

$Xi \sim [some\ distribution, with \mu, \sigma^2] \ \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
$Z_n = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{(\sum_{i=1}^n X_i) - n\mu}{\sqrt{n}\sigma}$ converges to standard Z as $n \to \infty$

6.3 Bivariate RV

$F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u,v) dv du$ and
$F_X(x) = F_{X,Y}(x, \infty), F_Y(y) = F_{X,Y}(\infty, y)$

$$P((x,y) \in A) = \iint_{(x,y) \in A} f_{X,Y}(x,y) dy dx$$

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

6.3.1 Bivariate Gaussian

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right)$$

6.4 Continuity Correction

for bernoulli RV Y, ensure the boundary equality holds before correction

$P(7 < X < 11) = P(8 \leq X \leq 10) = P(7.5 \leq X \leq 10.5)$

6.5 Complex Gaussian

$X \sim N(\mu_X, {\sigma_X}^2, Y \sim N(\mu_Y, {\sigma_Y}^2), H = X + iY$

$E[H] = \mu_X + i\mu_Y := \mu(complex)$

var $\Gamma = E[|H - E[H]|^2] = {\sigma_X}^2 + {\sigma_Y}^2 := \sigma^2$

6.5.1 Circularly Symmetric Gaussian RV

$X \sim N(0, \sigma^2/2, Y \sim N(0, \sigma^2/2), H = X + iY \sim CN(0, \sigma^2) = |H| exp(i\angle H) = R \times exp(i\theta)$ where
$R = |H| = \sqrt{X^2 + Y^2}, \theta = \angle H = \arctan(\frac{Y}{X}) \in [-\pi, \pi)$ (so that $(X,Y) \leftrightarrow (r, \theta)$ is bijective

$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} exp(-\frac{x^2+y^2}{2\sigma^2})$
$P(A) = \int_A f_{X,Y}(x,y) dx dy = \int_A f_H(r,\theta) dr d\theta$ (note that r is inside $f_H(r,\theta)$)

$f_H(r, \theta + \phi) = f_H(r, \theta) = f_{X,Y}(x,y) r = \frac{r}{2\pi\sigma^2} \exp(-\frac{r^2}{2\sigma^2})$

$f_R(r) = \int_{-\pi}^{\pi} f_H(r,\theta) d\theta = \frac{r}{\sigma^2} \exp(-\frac{r^2}{2\sigma^2}), r > 0$

$f_\theta(\theta) = \frac{1}{2\pi}, -\pi \leq \theta < \pi$

7 problem solving techniques

7.1 Mapping of Random Variables

7.2 Derived Distribution

1. $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{\{x | g(x) \leq y\}} f_X(x) dx$
2. $f_Y(y) = \frac{dF_Y(y)}{dy}$

If g is strictly monotonic and
$\exists h \ s.t. \forall x \in S_X, y = g(x) \Leftrightarrow x = h(y) = g^{-1}(y)$

$f_Y(y) = f_X(h(y)) |\frac{dh(y)}{dy}| > 0$

7.3 Graphical Methods

area as probability

Wong Jian Bin, National University of Singapore wongjianbin@u.nus.edu