15.085 Final Exam Cheat Sheet
by S. Shreyas V., page 1 of 2

1 Prob. Models & Measures

Prob. Experiments

Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$
–Ω sample space
–$\mathcal{F}$ σ-field
–$\mathbb{P}$ prob. measure

Disc. Prob. Space

Ω finite or countable,
$\mathcal{F}$ set of all subsets of Ω
$\mathbb{P} : \Omega \to [0,1]$ sums to 1

σ-fields

(a) $\emptyset \in \mathcal{F}$
(b) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
(c) $\{A_i\} \subseteq \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$
Set $A \in \mathcal{F}$ "event"/"measurable,"
$(\Omega, \mathcal{F})$ *measurable space*
$\mathcal{F} = \bigcap_{s \in S} \mathcal{F}_s$ also σ-field

Prob. Measures

Measure: $\mu : \mathcal{F} \to [0, \infty]$
(a) $\mu(\emptyset) = 0$;
(b) Countable additivity: $\{A_i\} \subseteq \mathcal{F}$ disjoint
$\implies \mu(\cup_i A_i) = \sum_{i=1}^{\infty} \mu(A_i)$
Prob. measure also has $\mathbb{P}(\Omega) = 1$
Field: like σ-field but finite

Continuity

Sequence of sets converge to union/intersection

2 Fundamental Models

Carathéodory's extension thm.

$\mathcal{F}_0$ field, $\mathcal{F}$ σ-field
$\mathbb{P}_0 : \mathcal{F}_0 \to [0,1]$, $\mathbb{P}_0(\Omega) = 1$
$\mathbb{P}_0$ yields $\mathbb{P}$ on $(\Omega, \mathcal{F})$

Lebesgue measure

Uniform measure on $[0,1]$

Borel σ-field

$\mathcal{B}$: smallest σ-field including every interval $[a,b] \subset [0,1]$
$A \subset [0,1]$, $A \in \mathcal{B}$ Borel set

3 Conditioning & Independence

Conditional probability

$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$
$\mathbb{P}_B(A) = \mathbb{P}(A \mid B)$
Bayes' rule: $\mathbb{P}_A(B_i) = \frac{\mathbb{P}(B_i)\mathbb{P}_{B_i}(A)}{\mathbb{P}(A)}$
$\mathbb{P}\left(\cap_{i=1}^{\infty} A_i\right) = \mathbb{P}(A_1) \prod_{i=2}^{\infty} \mathbb{P}\left(A_i \mid \cap_{j=1}^{i-1} A_j\right)$

Independence

Defn.: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
$\mathcal{F}_1, \mathcal{F}_2$ σ-fields indep. iff any $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$ indep.

Borel-Cantelli lemma

Sequence of events $\{A_n\}$,
$A = \{A_n \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$
(a) $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \implies \mathbb{P}(A) = 0$
(b) $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \implies \mathbb{P}(A) = 1$ if events A_n, $n \in \mathbb{N}$ indep.

Summation lemma

If $0 \le p_i \le 1$, $\forall i \in \mathbb{N}$, and $\sum_{i=1}^{\infty} p_i = \infty$, then $\prod_{i=1}^{\infty}(1 - p_i) = 0$

4 Combinatorial prob.

e

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

Stirling's approx.

$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

5 Random Variables

Definition

(a) $X : \Omega \to \mathbb{R}$ s.t. $\{\omega \mid X(\omega) \le c\}$ $\mathcal{F}$-measurable $\forall c \in \mathbb{R}$
(b) Extended-valued r.v. if $\forall c \in \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$
$X^{-1}(B) = \{X \in B\} = \{\omega \mid X(\omega) \in B\}$
Prob. law: $\mathbb{P}_X : \mathcal{B} \to [0,1]$, $B \mapsto \mathbb{P}(X \in B)$ (measure on $(\mathbb{R}, \mathcal{B})$)
$(\mathcal{F}_1, \mathcal{F}_2)$-measurable func.: $f : \Omega_1 \to \Omega_2$ s.t. $f^{-1}(B) \in \mathcal{F}_1, \forall B \in \mathcal{F}_2$
For $A \in \mathcal{F}$, I_A is $(\mathcal{F}, \mathcal{B})$-measurable

CDFs

Defn.: $F_X : \mathbb{R} \to [0,1]$, $x \mapsto \mathbb{P}(X \le x)$
(a) Monotonicity
(b) $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$
(c) Right-continuity

Discrete RV's

Range $X(\Omega)$ finite/countable
$p_X : \mathbb{R} \to [0,1]$, $x \mapsto \mathbb{P}(X = x)$ PMF

Continuous RV's

$F_X(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$, f PDF

6 Discrete RV's

Examples

Uniform: $p_X(k) = \frac{1}{b-a+1}$
Bernoulli: $p_X(1) = p$, $p_X(0) = 1 - p$
Binomial: $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$
Geometric: $p_X(k) = (1-p)^{k-1} p$
Poisson: $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$
Power law: $p_X(k) = \frac{1}{k^{\alpha}} - \frac{1}{(k+1)^{\alpha}}$

7 More Discrete RV's

Expected Values

Bernoulli: $\mathbb{E}[X] = p$, $\mathrm{Var}(X) = p(1-p)$
Binomial: $\mathbb{E}[X] = np$, $\mathrm{Var}(X) = np(1-p)$
Geometric: $\mathbb{E}[X] = \frac{1}{p}$, $\mathrm{Var}(X) = \frac{1-p}{p^2}$
Poisson: $\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$
Power law: $\mathbb{E}[X] = \sum_{k=0}^{\infty} \frac{1}{(k+1)^{\alpha}}$

Cov. & Corr.

$\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$
$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$ Cauchy-Schwarz ineq: $(\mathbb{E}[XY])^2 \le \mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]$

Conditional Expectations

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

8 Continuous RV's

Examples

Uniform: $F_X(x) = \frac{x-a}{b-a}$
Exponential: $F_X(x) = 1 - e^{-\lambda x}$
Normal: $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Cauchy: $f_X(x) = \frac{1}{\pi(1+x^2)}$
Power law: $f_X(t) = \frac{\alpha c^{\alpha}}{t^{\alpha+1}}$

Exp. Value

$\mathbb{E}[X] = \int_0^{\infty} (1 - F_X(t))\mathrm{d}t$ for X non-negative

Joint Dist.'s

$F_{X,Y} = \mathbb{P}(X \le x, Y \le y)$
$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$

Independence

$F_{X,Y} = F_X(x) F_Y(y)$

9 More Cont. RV's

Conditional PDFs

$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

Bivariate Normal Dist.

f_X, f_Y Normal PDF's

Mixed Bayes' Rule

Sums for discrete RV, prods for cont. RV

10 Derived Dist.'s

Func. of Single RV

If $Y = g(X)$, calculate $F_Y(y) = \mathbb{P}(g(X) \le y) = \int_{\{x | g(x) \le y\}} f_X(x)\mathrm{d}x$
Then $f_Y(y) = \frac{\mathrm{d}F_Y}{\mathrm{d}y}(y)$

Multivariate

$f_Y(y) = f_X\left(M^{-1}y\right) \cdot \left|M^{-1}\right|$

Max & Min of RV's

$\mathbb{P}\left(\max_j X_j \le x\right) = F_{X_1}(x) \cdots F_{X_n}(x)$
$\mathbb{P}\left(\min_j X_j \le x\right) = 1 - \left(1 - F_{X_1}(x)\right) \cdots \left(1 - F_{X_n}(x)\right)$

Convolution

$p_{X+Y}(z) = \sum_x p_X(x) p_Y(z - x)$
$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\mathrm{d}x$

11 Abstract Integration I

Preliminaries

$(\Sigma, \mathcal{F}, \mathbb{P})$: $\mathbb{E}[X] = \int X \mathrm{d}\mathbb{P}$
$(\mathbb{R}, \mathcal{B}, \lambda)$: $\int g \mathrm{d}\lambda = \int g(x)\mathrm{d}x$

Results

Monotone Convergence Theorem:
$0 \le g_n \uparrow g \implies \int g_n \mathrm{d}\mu \uparrow \int g \mathrm{d}u$

General Case

Let $g_+ = g \cdot \mathbb{1}_{g>0}$, $g_- = -g \cdot \mathbb{1}_{g<0}$:
$\int g \mathrm{d}\mu = \int g_+ \mathrm{d}\mu - \int g_- \mathrm{d}\mu$

12 Abstract Integration II

Fatou's Lemma

Y s.t. $\mathbb{E}[|Y|] < \infty$: (a) If $Y \le X_n \forall n$, then $\mathbb{E}[\liminf_{n \to \infty} X_n] \le \liminf_{n \to \infty} \mathbb{E}[X_n]$
(b) If $X_n \le Y \forall n$, then $\mathbb{E}[\limsup_{n \to \infty} X_n] \ge \limsup_{n \to \infty} \mathbb{E}[X_n]$

Dominated Convergence Theorem

Sequence $\{X_n\}$ converges to X a.e.
Suppose $|X_n| \le Y$, $\forall n$,
$Y \ge 0$ with $\mathbb{E}[Y] < \infty$.
Then: $\lim_{n \to \infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.
Corollary: if $\sum_{n=1}^{\infty} \mathbb{E}[|Z_n|] < \infty$, $\sum_{n=1}^{\infty} \mathbb{E}[Z_n] = \mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right]$.

13 Product Measure & Fubini's Thm

Product Measure

$\mathcal{F}_1 \times \mathcal{F}_2$: smallest σ-field of subsets of $\Omega_1 \times \Omega_2$ containing all $A_1 \times \Omega_2$ and $\Omega_1 \times A_2$, for $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$. Exists unique $\mathbb{P}$ on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ with $\mathbb{P}(A_1 \times A_2) = \mathbb{P}_1(A_1)\mathbb{P}_2(A_2)$

Fubini's Theorem

If g nonnegative or $\int_{\Omega_1 \times \Omega_2} |g(\omega_1, \omega_2)| \mathrm{d}\mathbb{P} < \infty$:
can switch order of integration

14 Moment Generating Functions

MGF's

$M_X(s) = \mathbb{E}[\exp(sX)]$
Domain $D_x = \{s \mid M_x(s) < \infty\}$
Inversion Thm: if $M_X(s) = M_Y(s)$, $\forall |s| \le a$, then $F_X = F_Y$

Probability GF: $g_X(s) = \mathbb{E}\left[s^X\right]$
$\frac{\mathrm{d}}{\mathrm{d}s} g_X(s)\Big|_{s=1} = \mathbb{E}[X]$
$Y = aX + b \implies M_Y(s) = e^{sb} M_X(as)$
X and Y indep.: $M_{X+Y}(s) = M_X(s) M_Y(s)$
If $\mathbb{P}(Z = X) = p$ and $\mathbb{P}(Z = Y) = 1 - p$,
$M_Z(s) = p M_X(s) + (1-p) M_Y(s)$

Sum of RVs

Law of Total Variance: $\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(\mathbb{E}[Y \mid X])$

15 Multivariate Normal Dist.

Positive Definite

A is $n \times n$ symmetric matrix
$A > 0 \iff x^{\intercal} A x > 0$, $\forall x \in \mathbb{R}^n$,
$A \ge 0 \iff x^{\intercal} A x \ge 0$, $\forall x \in \mathbb{R}^n$.
Symmetric: n real eigenvalues; Pos Defn.: n real positive eigenvalues;
Nonnegative Defn.: n real nonnegative eigenvalues.
Symmetric matrices are diagonalizable

Multivariate Normal Dist.

Defn. of mv normal:
$f_X(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |V|}} \exp\left\{-\frac{(\mathbf{x}-\mu)V^{-1}(\mathbf{x}-\mu)^{\intercal}}{2}\right\}$;
$\mathbf{X} = D\mathbf{W} + \mu$, $W_{ij} \sim \mathcal{N}(0,1)$;
$\forall \mathbf{a} \in \mathbb{R}^n$: $\mathbf{a}^{\intercal}\mathbf{X} \sim \mathcal{N}\left(\mu, \sigma^2\right)$
Factorization: convert X_i to W_i
$W_1 = X_1$,
$W_2 = X_2 - \mathbb{E}[X_2 \mid X_1]$,
$W_n = X_n - \mathbb{E}[X_n \mid X_1, \ldots, X_{n-1}]$

16 Characteristic Functions

Basics

$\phi_X(t) = \mathbb{E}\left[e^{itX}\right]$
Can use $e^{ix} = \cos(x) + i\sin(x)$
Invertible for all RV's:
$f_X(x) = \frac{1}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} e^{-itx} \phi_X(t)\mathrm{d}x$
$\lim_{n \to \infty} X_n = X \implies \lim_{n \to \infty} \phi_{X_n}(t) = \phi_X(t)$
$\mathbb{E}\left[|X|^k\right] < \infty \implies \frac{\mathrm{d}^k}{\mathrm{d}t^k} \phi_X(t)\Big|_{t=0} = i^k \mathbb{E}\left[X^k\right]$
Exponential: $\phi_X(t) = \frac{\lambda}{\lambda - it}$
Normal: $\phi_X(t) = \exp\left(it\mu - \frac{t^2\sigma^2}{2}\right)$

17 MGF Applications

Random Walks

Can apply MGFs to calculate probability and time to return

15.085 Final Exam Cheat Sheet
by S. Shreyas V., page 2 of 2

Branching Processes

Z_n: indivs in n^{th} gen,

g is PGF of dist. F, $X \stackrel{d}{=} F$, $g(s) = \sum_{m\geq 0} s^m \mathbb{P}(X = m)$.

G_n is PGF of Z_n, $\mathbb{E}[X] = \mu$, $\text{Var}(X) = \sigma^2$, $G_n = g^{(n)}$, $\forall n \geq 0$.

$\mathbb{E}[Z_n] = \mu^n$,

$\text{Var}(Z_n) = \begin{cases} n\sigma^2, & \mu = 1; \\ \frac{\sigma^2(\mu^n - 1)\mu^{n-1}}{\mu - 1} & \mu \neq 1. \end{cases}$

If $\mu < 1$, $Z_n \to 0$, $\mathbb{P}(Z_n > 0) < \mu^n$. Pr. extinction η smallest root of $s = g(s)$; $\eta = 1$ in many cases.

18 Convergence

Definition

Almost surely: $X_n \xrightarrow{\text{a.s.}} X$ if $\exists A \subset \Omega$:
(a) $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, $\forall \omega \in A$,
(b) $\mathbb{P}(A) = 1$.

Dist: $X_n \xrightarrow{\text{d}} X$ if (F and F_n CDFs) $\lim_{n\to\infty} F_n(x) = F(x)$, $\forall x \in \mathbb{R}$.

Prob: $X_n \xrightarrow{\text{i.p.}} X$ if ($\forall \varepsilon > 0$) $\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$.

Hierarchy

Each implies the next:
i) Almost Surely
ii) In Probability
ii) In Distribution
iv) $\phi_{X_n}(t) \to \phi_X(t), \forall t$
Last one holds in reverse

19 LLN & CLT

Inequalities

Markov: $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$
Chebyshev: extends Markov $\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}$

Weak LLN

X_n i.i.d., $\mathbb{E}[|X_1|] < \infty$, S_n sum
$\frac{S_n}{n} \xrightarrow{\text{i.p.}} \mathbb{E}[X_1]$; Strong LLN a.s.

CLT

X_n i.i.d., mean μ, var σ^2, S_n sum
$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\text{d}} X \sim \mathcal{N}(0,1)$

20 LLN II

Strong LLN

$\{X_n\}$ sequence of RV's
(i) $\sum_{n=1}^{\infty} \mathbb{E}[|X_n|^s] < \infty$, $s > 0 \implies X_n \xrightarrow{\text{a.s.}} 0$
(ii) $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \varepsilon) < \infty$, $\forall \varepsilon > 0 \implies X_n \xrightarrow{\text{a.s.}} 0$
LLN: $X, X_1, X_2, \ldots$ i.i.d., $\mathbb{E}[|X|] < \infty$, S_n sum $\implies S_n/n \xrightarrow{\text{a.s.}} \mathbb{E}[X]$.

Chernoff Bound

If $\mathbb{E}[\exp(sX)] < \infty$ for some $s > 0$, and $a > 0$, then $\mathbb{P}(S_n \geq na) \leq \exp(-n\phi(a))$, where $\phi(a) = \sup_{s\geq 0}(sa - \ln M_X(s))$.
If X continuous and boundless, $\lim_{n\to\infty} n^{-1}\mathbb{P}(S_n \geq na) = -\phi(a)$.

21 Stochastic Processes

Bernoulli

$X_n \sim \text{Ber}(p)$ i.i.d., $\Omega = \{0,1\}^{\infty}$.
S_n: $\mathbb{E}[S_n] = np$, $\text{Var}(S_n) = npq$.
$T_1 = \min\{n \mid X_n = 1\}$, $\mathbb{E}[T_1] = p^{-1}$.
Stationary and memoryless

Poisson

$N(t)$: arrivals during $(0,t]$,
$N(t) \sim \text{Pois}(\lambda t)$.
$\mathbb{P}(T_1 > t) = \exp(-\lambda t)$
Given info on number of arrivals by t, prev. arrivals are dist. $\mathcal{U}(0,t]$.

22 Markov Chains

Basics

Process takes values in countable $\mathcal{X}$, probability $X_n = x_n$ only conditioned on $X_{n-1} = x_{n-1}$
Homogeneous: constant transition matrix $\mathbf{P}$;
$p_{i,j} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$ s.t.
$\sum_j p_{i,j} = 1$, $\forall i$—stochastic matrix

Stationary Distribution

π stationary iff $\pi^\intercal = \pi^\intercal \mathbf{P}$
Every finite state Markov chain has at least 1 stationary distribution

State Classification

Transient: i s.t. $\exists j : i \to j, j \not\to i$.
Recurrent: i not transient.

23 Markov Chains II

Single Recurrence Class

$\leftrightarrow$ is an equivalency relation: recurrent states form classes $R_1, \ldots, R_r$. $\mathcal{T}$ transient states.
$\forall l = 1, \ldots, r$, $\forall i \in R_l, j \notin R_l$: $p_{i,j} = 0$
Let $T_i = \min\{n \geq 1 : X_n = i \mid X_0 = i\}$ first passage time; $\mu_i = \mathbb{E}[T_i]$ mean recurrence time.
$\forall i \in \mathcal{T} : \mathbb{P}(X_n = i, \text{i.o.}) = 0$, $\mu_i = \infty$.
Irreducible: $\mathcal{X} = \mathcal{T} \cup R$ and $\mathcal{T} = \emptyset$.

Uniqueness of π

If single recurrence class, π unique.
Furthermore, $\pi_i = \mu_i^{-1}$, $\forall i$.
If i recurrent, $\pi_i > 0$.

Ergodic Theorem

Let $N_i(t)$ be # of times state i visited during times $0, 1, \ldots t$.
For arbitrary starting $X_0 = k$, $\forall i$,
$\lim_{t\to\infty} \frac{N_i(t)}{t} = \pi_i$ a.s.
$\lim_{t\to\infty} \frac{\mathbb{E}[N_i(t)]}{t} = \pi_i$.

Multiple Recurrence Classes

r stationary distributions where $\forall 1 \leq i \leq r$: $\pi_k^i = 0$ if $k \notin R_i$ and $\pi_k^i = \mu_k^{-1}$ if $k \in R_i$.

24 Markov Chains III

Periodicity

$\forall x \in \bigcup R_i$, $I_x = \left\{n \geq 1 : p_{x,x}^{(n)} > 0\right\}$
d_x period is GCD of I_x numbers
States in same recurrence class have same period
Periodic: $d_x > 1$;
Aperiodic: $d_x = 1$
$d_y = 1 \implies \exists N \geq 1 : p_{y,y}^{(n)} > 0, \forall n \geq N$.

Coupling & Mixing

Irreducible aperiodic MC: $\forall x, y \in \mathcal{X} : \lim_{n\to\infty} p_{x,y}^{(n)} = \pi_y$.
Irreducible aperiodic reversible MC: $\exists C : \forall x, y \in \mathcal{X} : \left|p_{x,y}^{(n)} - \pi_y\right| \leq C|\lambda_2|^n$, where λ_2 second largest (absolute value) eigenvalue of $\mathbf{P}$.
Coupling: X_n MC on $\{1, \ldots, N\}$ and Y_n MC on $\{1, \ldots, M\}$ with $\mathbf{P}$, $\mathbf{Q}$, respectively. Can create Z_n MC on $\{1, \ldots, N\} \times \{1, \ldots, M\}$ with transition $R = \left\{r_{(x_1,x_2),(y_1,y_2)}\right\}$.
Moves in each direction with same probability matrices.
Coupling behaves as above until collision—then transition the same.

Absorbtion Probability

Assume recurrent states i absorbing, i.e., $p_{i,i} = 1$.
Absorbing probability: $a_{ki} = \mathbb{P}(X_n = i \text{ eventually} \mid X_0 = k)$.