EECS 126 Midterm 1 Sheet Wang Jianzhi

Known Distributions

Bernoulli Distribution: $X \sim \text{Bernoulli}(p)$

$\mathbb{P}[X=0]=1-p$ $\quad \mathbb{E}[X]=p$ $\quad F(x)=\begin{cases}0 & x<0\\ 1-p & 0\le x<1\\ 1 & x\ge 1\end{cases}$

$\mathbb{P}[X=1]=p$ $\quad Var[X]=p(1-p)$

Binomial Distribution: $X\sim \text{Binomial}(n,p)$

$\mathbb{P}[X=i]=\binom{n}{i}p^i(1-p)^{n-i}$, $i=0,1,2,\dots n$

$\mathbb{E}[X]=np$ $\quad Var[X]=np(1-p)$

Geometric Distribution: $X\sim \text{Geom}(p)$

$\mathbb{P}[X=i]=(1-p)^{i-1}p$, $i=1,2,\dots$ $\quad \mathbb{E}[X]=\frac{1}{p}$ $\quad Var[X]=\frac{1-p}{p^2}$

$pgf(x)=\frac{px}{1-(1-p)x}$; $F(x)=1-(1-p)^x$; $\mathbb{P}[X>n+m\mid X>n]=\mathbb{P}[X>m]$ $\quad n,m>0$

Poisson Distribution: $X\sim\text{Poisson}(\lambda)$

$\mathbb{P}[X=i]=\frac{\lambda^i e^{-\lambda}}{i!}$, $i=0,1,2,\dots$ $\quad \mathbb{E}[X]=\lambda$, $Var[X]=\lambda$

$X+Y\sim \text{Poisson}(\lambda+\mu)$; $\text{Binomial}(n,\frac{\lambda}{n})\xrightarrow{n\to\infty}\text{Poisson}(\lambda)$

Exponential Distribution: $X\sim\text{Expo}(\lambda)$

$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0 & \text{otherwise}\end{cases}$ $\quad \mathbb{E}[X]=\frac{1}{\lambda}$, $Var[X]=\frac{1}{\lambda^2}$, $\mathbb{E}[X^2]=\frac{2}{\lambda^2}$, $F_X(x)=1-e^{-\lambda x}$

$X\sim\text{Expo}(\lambda)$, $Y\sim\text{Expo}(\mu)\Rightarrow \min(X,Y)\sim\text{Expo}(\lambda+\mu)$

$X_i\sim\text{Expo}(\lambda_i)\Rightarrow \mathbb{P}(X_i=\min(X_i))=\frac{\lambda_i}{\sum\lambda_i}$

Discrete Uniform: $X\sim\text{Uniform}(\{a,\dots,b\})$ $\quad P_X(k)=\begin{cases}\frac{1}{b-a+1}\\ 0 & \text{else}\end{cases}$

$\mathbb{E}[X]=\frac{a+b}{2}$ $\quad Var(X)=\frac{(b-a)(b-a+2)}{12}$

Continuous Uniform $X\sim\text{Uniform}([a,b])$

$\mathbb{E}[X]=\frac{a+b}{2}$; $Var[X]=\frac{(b-a)^2}{12}$; $f_X(x)=\frac{1}{b-a}$; $F_X(x)=\frac{x-a}{b-a}$.

Normal Distribution: $X\sim N(\mu,\sigma^2)$ $\quad Z\sim N(0,1)$

$f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$; $f_Z(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$; $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$

$X\sim N(0,\sigma^2)\Rightarrow \mathbb{E}[X^{2n}]=(2n-1)!!\,\sigma^{2n}$.

Erlang Distribution: $X\sim \text{Erlang}(k,\lambda)$ (Sum of k iid $\text{Expo}(\lambda)$)

$\mathbb{E}[X]=\frac{k}{\lambda}$. $Var[X]=\frac{k}{\lambda^2}$; $f_X(x)=\frac{\lambda^k x^{k-1}e^{-\lambda x}}{(k-1)!}$; $X\sim\text{Erlang}(k,\lambda)\Rightarrow aX\sim\text{Erlang}(k,\frac{\lambda}{a})$

$F_X(x)=1-\sum_{n=0}^{k-1}\frac{1}{n!}e^{-\lambda x}(\lambda x)^n$; $X\sim\text{Erlang}(k_1,\lambda)$ $Y\sim\text{Erlang}(k_2,\lambda)$ $\Rightarrow X+Y\sim\text{Erlang}(k_1+k_2,\lambda)$

Pascal Distribution $X\sim\text{Pascal}(k,p)$ (Sum of k iid $\text{Geom}(p)$)

$\mathbb{E}[X]=\frac{k}{p}$; $Var[X]=\frac{k(1-p)}{p^2}$; $\mathbb{E}[X^2]=\frac{k^2+k(1-p)}{p^2}$

$P_X(x)=\binom{x-1}{k-1}p^k(1-p)^{x-k}$, $k=1,2,\dots$; $x=k,k+1,\dots$

$F_X(x)=\sum_{n=k}^{x}\binom{n-1}{k-1}p^k(1-p)^{n-k}$

Joint and Conditional Probabilities

- $P_{X|Y}(x|y)=\frac{\mathbb{P}[X=x,Y=y]}{\mathbb{P}[Y=y]}=\frac{P_{X,Y}(x,y)}{P_Y(y)}$
- $P_{X,Y}(x,y)=P_Y(y)\,P_{X|Y}(x|y)$
- $f_{X,Y}(x,y)=\frac{\partial F_{X,Y}}{\partial x\partial y}(x,y)$ $\quad$ • $f_{X,Y}(x,y)=f_Y(y)f_{X|Y}(x|y)$
- $f_X(x)=\int_{-\infty}^{\infty}f_Y(y)f_{X|Y}(x|y)\,dy$
- $f_X(x)=\sum_{i=1}^{n}f_{X|A_i}(x)\,\mathbb{P}[A_i]$

Convolution

$\mathbb{P}[Z=z]=\mathbb{P}[X+Y=z]=\sum_x\mathbb{P}[X=x]\,\mathbb{P}[Y=z-x]$

$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\,dx$.

Probabilistic Bounding

1. [Markov] nonnegative X, finite mean $\mathbb{P}[X\ge c]\le\frac{\mathbb{E}[X]}{c}$, $c>0$.
2. [Chebyshev] $c>0$. $\mathbb{P}[|X-\mu|\ge c]\le\frac{Var[X]}{c^2}$ $\quad \mathbb{P}[|X-\mu|\ge k\sigma]\le\frac{1}{k^2}$.
3. [Generalized Markov] Y not necessarily nonnegative finite mean; $c,r>0$ $\quad\mathbb{P}[|Y|\ge c]\le\frac{\mathbb{E}[|Y|^r]}{c^r}$
4. [Extended Markov] X not necessarily nonnegative, $\Phi(x)$ nonnegative function, monotonically increasing for $x>0$, $\alpha>0$ $\quad\mathbb{P}[X\ge\alpha]\le\frac{\mathbb{E}[\Phi(X)]}{\Phi(\alpha)}$
5. [Cantelli] $\alpha>0$. $\mathbb{P}[X-\mathbb{E}[X]\ge\alpha]\le\frac{\sigma^2}{\alpha^2+\sigma^2}$
6. [Law of Large Numbers] $X_1,\dots,X_n$ i.i.d. RV with $\mu<\infty$ $\quad S_n=X_1+\dots+X_n$. $\forall\varepsilon$, $\lim_{n\to\infty}\mathbb{P}[|\frac{1}{n}S_n-\mu|<\varepsilon]=1$
7. [Central Limit Theorem] Distribution of $\frac{M_n-\mu}{\sqrt{\sigma^2/n}}$ approaches a normal distribution
   $\mathbb{P}[\frac{S_n-n\mu}{\sigma\sqrt{n}}\le c]\longrightarrow\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c}e^{-\frac{x^2}{2}}dx$
   i.e. define $Z_n=\frac{X_1+\dots+X_n-n\mu}{\sigma\sqrt{n}}$, then $(Z_n)_n\xrightarrow{d}N(0,1)$ in distribution.
8. [Chernoff] $\mathbb{P}[X\ge a]=\mathbb{P}[e^{sX}\ge e^{sa}]\le\frac{M_X(s)}{e^{sa}}$ $\forall s\ge 0$
9. [Jensen] $f(\mathbb{E}[X])\le\mathbb{E}[f(X)]$ $\forall f$ convex.
10. [WLLN] Let $(X_n)_n$ be a sequence of i.i.d. RV with mean μ. $\forall\varepsilon>0$, $\mathbb{P}[|\frac{X_1+\dots+X_n}{n}-\mu|\ge\varepsilon]\longrightarrow 0$.
11. [SLLN] Let $(X_n)_n$ be a sequence of i.i.d. RV with mean μ. Then $M_n=\frac{X_1+\dots+X_n}{n}$ converges to μ with probability 1
    $\mathbb{P}[\lim_{n\to\infty}\frac{X_1+\dots+X_n}{n}=\mu]=1$. (i.e. $\forall\varepsilon>0$, $\mathbb{P}(|M_n-\mu|>\varepsilon \text{ i.o.})=0$)
12. [De Moivre-Laplace Approximation] If $S_n\sim\text{Binom}(n,p)$ and $n\gg1$, k,l nonnegative, then
    $\mathbb{P}[k\le S_n\le l]\approx\left[\Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)\right]$

Conditional Expectation and Variance

Note: $\mathbb{E}[Y|X]=f(X)$ i.e. is a function of X

$\mathbb{E}[Y|X=x]$ is a real number

To find $\mathbb{E}[Y|X]$, generalize pattern from $\mathbb{E}[Y|X=x]$

1. (Linearity) $\mathbb{E}[a_1Y_1+a_2Y_2|X]=a_1\mathbb{E}[Y_1|X]+a_2\mathbb{E}[Y_2|X]$
2. (Factoring Known Values) $\mathbb{E}[h(X)Y|X]=h(X)\mathbb{E}[Y|X]$
3. (Independence) $\mathbb{E}[X|Y]=\mathbb{E}[X]$ (X,Y independent)

$\mathbb{E}[X]=\mathbb{E}[\mathbb{E}[X|Y]]$; $\mathbb{E}[XY]=\mathbb{E}[\mathbb{E}[XY|Y]]=\mathbb{E}[Y\mathbb{E}[X|Y]]$

$Var[X]=\mathbb{E}[Var[X|Y]]+Var[\mathbb{E}[X|Y]]$

Conditioning on Event

- $\mathbb{E}[X|A]=\int_{-\infty}^{\infty}x f_{X|A}(x)\,dx$
- $\mathbb{E}[g(X)|A]=\int_{-\infty}^{\infty}g(x)f_{X|A}(x)\,dx$
- $\mathbb{E}[X]=\sum_{i=1}^{n}\mathbb{P}(A_i)\mathbb{E}[X|A_i]$

Conditioning on RV

- $\mathbb{E}[Y]=\mathbb{E}[\mathbb{E}[Y|X]]=\int_0^{\infty}\mathbb{E}[Y|X=x]f_X(x)\,dx$
- $\mathbb{E}[X|Y=y]=\int_{-\infty}^{\infty}x f_{X|Y}(x|y)\,dx$
- $\mathbb{E}[g(X)|Y=y]=\int_{-\infty}^{\infty}g(x)f_{X|Y}(x|y)\,dx$
- $\mathbb{E}[g(X,Y)|Y=y]=\int_{-\infty}^{\infty}g(x,y)f_{X|Y}(x|y)\,dx$

Bayes and Continuous Bayes

$\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B]}$ $f_Y(y) f_{X|Y}(x|y) = f_X(x) f_{Y|X}(y|x)$

- $f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(t) f_{Y|X}(y|t)\,dt}$
- $f_Y(y)\,\mathbb{P}[N=n|Y=y] = \mathbb{P}[N=n]\, f_{Y|N}(y|n)$
- $\mathbb{P}[N=n|Y=y] = \frac{\mathbb{P}[N=n]\, f_{Y|N}(y|n)}{\sum_i \mathbb{P}[N=i]\, f_{Y|N}(y|i)}$
- $f_{Y|N}(y|n) = \frac{f_Y(y)\,\mathbb{P}[N=n|Y=y]}{\int_{-\infty}^{\infty} f_Y(t)\,\mathbb{P}[N=n|Y=t]\,dt}$

Differential Probability

$f_X(x)dx = \mathbb{P}[x \le X \le x+dx]$; $f_{X|Y}(x|y)dx = \frac{\mathbb{P}[x\le X\le x+dx,\ y\le Y\le y+dy]}{\mathbb{P}[y\le Y\le y+dy]}$

- $\frac{d}{dz}\mathbb{P}[Z\le z|X=x] = f_{Z|X}(z|x)$
- $f_Y(y) = f_X(f^{-1}(y))\left|\frac{df^{-1}}{dy}(y)\right|$

Moment Generating Functions (MGF)

Properties: $e^{sX} = 1 + \frac{sX}{1!} + \frac{(sX)^2}{2!} + \frac{(sX)^3}{3!} + \dots$

$\mathbb{E}[e^{sX}] = M_X(s) = 1 + s\mathbb{E}[X] + \frac{s^2}{2!}\mathbb{E}[X^2] + \frac{s^3}{3!}\mathbb{E}[X^3] + \dots$

$\left(\frac{d^n}{ds^n}\mathbb{E}[e^{sX}]\right)\Big|_{s=0} = \mathbb{E}[X^n]$; $M_X(0)=1$ independent

$Y = aX+b \Rightarrow M_Y(s) = e^{sb}M_X(as)$; $Z = \sum X_i \Rightarrow M_Z(s) = \prod M_{X_i}(s)$

Distributions:

Bernoulli(p): $M(s) = (1-p)+pe^s$; Binomial(n,p): $M(s) = (1-p+pe^s)^n$

Geometric(p): $M(s) = \frac{pe^s}{1-(1-p)e^s}$; Poisson(λ): $M(s) = e^{\lambda(e^s-1)}$

Expo(λ): $M(s) = \frac{\lambda}{\lambda-s}$ (λ>s); Uniform([a,b]): $M(s) = \frac{e^{sb}-e^{sa}}{s(b-a)}$

$N(\mu,\sigma^2)$: $M(s) = e^{\mu s+\frac{\sigma^2 s^2}{2}}$; Erlang(k,λ): $M(s) = \left(\frac{\lambda-s}{\lambda}\right)^k$

Pascal(k,p): $M(s) = \left(\frac{pe^s}{1-(1-p)e^s}\right)^k$

Convergence and Theorems

Convergence in Distribution: A sequence $(X_n)_n$ converges in distribution to X, denoted as $X_n \xrightarrow{d} X$ if $\forall x$, $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ (i.e. CDF of X_n converges to CDF X)

Theorem: For integer valued X, $(X_n)_{n\ge 1}$, suffices to show $\forall x$ $\mathbb{P}[X=x] = \lim_{n\to\infty}\mathbb{P}[X_n=x]$.

Convergence in Probability: A sequence $(X_n)_n$ converges in probability to X denoted as $X_n \xrightarrow{P} X$ if $\forall \varepsilon>0$, $\lim_{n\to\infty}\mathbb{P}[|X_n-X|\ge\varepsilon] = 0$

Theorem: If $X_n \xrightarrow{d} c$ constant, then $X_n \xrightarrow{P} c$.

Convergence with Probability 1: A sequence $(X_n)_n$ converges to X with probability 1, denoted as $X_n \xrightarrow{as} X$ if under sample space Ω, $\mathbb{P}[\{\omega\in\Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}] = 1$

Theorem: Consider $(X_n)_n$. If $\forall\varepsilon>0$ $\sum_{n=1}^{\infty}\mathbb{P}[|X_n-X|>\varepsilon]<\infty$, then $X_n \xrightarrow{as} X$

Theorem: Consider $(X_n)_n$. For ε>0, define $A_m = \{|X_n-X|<\varepsilon\ \forall n\ge m\}$ then $X_n \xrightarrow{as} X$ if and only if $\lim_{m\to\infty}\mathbb{P}[A_m] = 1$.

Miscellaneous

$f_{X_{(i)}}(x) = n\binom{n-1}{i-1} f(x) F(x)^{i-1}(1-F(x))^{n-i}$ $\mathbb{E}[X^{(i)}] = \frac{i}{n+1}$

$\mathbb{E}[X] = \sum_{i=1}^{\infty}\mathbb{P}[X\ge i]$ $\mathbb{E}[Z] = \int_0^{\infty}\mathbb{P}[Z\ge z]\,dz$

Conditional Expectation and Variance

Series of RV

- $Y = X_1+X_2+\dots+X_N$
- $\mathbb{E}[Y|N] = N\cdot\mathbb{E}[X]$
- $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|N]] = \mathbb{E}[N]\mathbb{E}[X]$
- $Var[Y] = \mathbb{E}[N]Var[X] + \mathbb{E}[X]^2 Var[N]$
- $M_Y(s) = \sum_{n=0}^{\infty}(M_X(s))^n\,\mathbb{P}[N=n]$

Conditioning on RV

- $\mathbb{E}[g(X,Y)] = \mathbb{E}[\mathbb{E}[g(X,Y)|Y]] = \int_{-\infty}^{\infty}\mathbb{E}[g(X,Y)|Y=y]f_Y(y)\,dy$
- $\mathbb{E}[X|N>k] = \mathbb{E}[\mathbb{E}[X|N]|N>k] = \sum_{n=1}^{\infty}\mathbb{E}[X|N=n, N>k]\,\mathbb{P}[N=n|N>k]$

Covariance and Correlation

Covariance: $Cov[X,Y] = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]$

$Cov[X,X] = Var[X]$; $Var(X+Y) = Var(X)+Var(Y)+2Cov(X,Y)$

$Cov[aX_1+bX_2, cY_1+dY_2] = ac\,Cov[X_1,Y_1]+ad\,Cov[X_1,Y_2] + bc\,Cov(X_2,Y_1)+bd\,Cov(X_2,Y_2)$

X,Y independent $\Rightarrow Cov(X,Y)=0$

Correlation: $\rho(X,Y) = \frac{Cov[X,Y]}{\sigma_X\sigma_Y}$; $X' = \frac{X-\mu_X}{\sigma_X}$; $Y' = \frac{Y-\mu_Y}{\sigma_Y}$

$-1 \le \rho[X,Y] = Cov(X',Y') \le 1$

$\rho[X,Y] = 1 \Rightarrow Y = AX+B,\ A>0\ (Y'=X')$

$\rho(X,Y) = -1 \Rightarrow Y = AX+B,\ A<0\ (Y'=-X')$

Borel Cantelli lemma and Continuity

First lemma: Let $(A_n)_n$ be sequence of events. If $\sum_{n=1}^{\infty}\mathbb{P}[A_n]<\infty$ then $\mathbb{P}[A_n\ i.o.] = 0$ (i.e. probability that infinitely many of A_n occurring is 0).

Second lemma: Let $(A_n)_n$ be sequence of events. If A_n independent and $\sum_{n=1}^{\infty}\mathbb{P}[A_n] = \infty$, then $\mathbb{P}[A_n\ i.o.] = 1$ (i.e. probability that infinitely many of them occurring is 1).

Let $h:\mathbb{R}\to\mathbb{R}$ be continuous:

If $X_n \xrightarrow{d} X$, then $h(X_n) \xrightarrow{d} h(X)$. If $X_n \xrightarrow{P} X$, $h(X_n) \xrightarrow{P} h(X)$.

If $X_n \xrightarrow{as} X$, then $h(X_n) \xrightarrow{as} h(X)$

Classics

- [Ballot] Let A,B be players s.t. A scored $n > m$ points and B scored m points. Then $\mathbb{P}$[A always ahead of B] $= \frac{n-m}{n+m}$.
- [Gambler's Ruin] Let A be a player who start at i and at each step increment by $\{-1,+1\}$ with probability $\frac{1}{2}$. Then $\mathbb{P}$[A reaches n] $= \frac{i}{n}$.
- [Gambler's Ruin (unfair)] If step probability is p, then $\mathbb{P}$[A reaches n] $= \frac{1-\left(\frac{p}{1-p}\right)^{n-i}}{1-\left(\frac{p}{1-p}\right)^{n}}$
- [Secretary] Optimal cutoff is ne^{-1}.
- [Coupon collector] $nH_n = n\sum_{i=1}^{n}\frac{1}{i}$.

Last Resort and Final Checks

- Indicators! $Var[X] = Cov[X,X]$; Symmetry Bijection
- Draw Pictures
- Define all ranges for f_X, F_X. Isolate objects
- One off error, conditioning, intuition, momentum.

$\mathbb{E}[X] = \int_0^{\infty}\mathbb{P}[X>x]\,dx - \int_0^{\infty}\mathbb{P}[X<-x]\,dx$; $\int_0^1 y^a(1-y)^b\,dy = \frac{a!b!}{(a+b+1)!}$

$\mathbb{P}[X=0] \le \frac{Var[X]}{\mathbb{E}[X]^2}$; $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n}\mathbb{P}(A_i)$; $X\sim N(0,1)$ $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$