Distributions

Bernoulli: $X \sim \text{Bernoulli}(p)$
$P(X=0)=p$ $E[X]=p$
$P(X=1)=(1-p)$ $Var(X)=p(1-p)$

Binomial: $X \sim Bin(n,p)$
$P(X=i)=\binom{n}{i}p^i(1-p)^{n-i}$ $i=0,1,\dots,n$
$E[X]=np$ $var(X)=np(1-p)$

Geometric: $X \sim geo(p)$
$P(X=i)=(1-p)^{i-1}p$ $i=1,2,\dots$
$E[X]=\frac{1}{p}$ $Var(X)=\frac{1-p}{p^2}$
$P(X \le x)=1-(1-p)^x$

Poisson: $X \sim poisson(\lambda)$
$P(X=i)=\frac{\lambda^i}{i!}e^{-\lambda}$, $i=0,1,\dots$
$E[X]=\lambda$ $var(X)=\lambda$
$X \sim poisson(\lambda)$ $Y \sim poisson(\mu)$
$(X+Y) \sim poisson(\mu+\lambda)$

Exponential: $X \sim exp(\lambda)$
$f(x)= \lambda e^{-\lambda x}$ $x \geq 0$; 0 otherwise
$F_X(x)=1-e^{-\lambda x}$
$E(X)=\frac{1}{\lambda}$ $var(X)=\frac{1}{\lambda^2}$
$X \sim exp(\lambda)$ $Y \sim exp(\mu)$
$\min(x,y) \sim exp(\lambda+\mu)$
$P(X \le Y)=\frac{\lambda}{\lambda+\mu}$

Discrete uniform $X \sim Uni(\{a,\dots,b\})$
$E[X]=\frac{a+b}{2}$ $Var(X)=\frac{(b-a)(b-a+2)}{12}$
$f(x)=\frac{1}{b-a}$ $F_X(x)=\frac{x-a}{b-a}$

Continuous uniform $X \sim uni([a,b])$
$E[X]=\frac{a+b}{2}$ $var(X)=\frac{(b-a)^2}{12}$
$f(x)=\frac{1}{b-a}$ $F_X(x)=\frac{x-a}{b-a}$

Normal: $X \sim N(\mu,\sigma^2)$
$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
$X \sim N(\mu,\sigma^2)$ $Z=\frac{X-\mu}{\sigma^2} \sim N(0,1)$
$f_Z(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$
$X \sim N(0,\sigma^2) \to E[X^{2n}]=\sigma^{2n}(2n-1)!!$
$X \sim N(\mu_X,\sigma_X^2)$ $Y \sim N(\mu_Y,\sigma_Y^2)$
$X+Y \sim (\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)$
$\alpha X+c \sim N(\alpha\mu_X+c, \alpha^2\sigma_X^2)$

Joint & conditional prob.

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

Bayes & conditional Bayes

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

Random Variable as fnc. of another

$Y \sim g(X)$ if g is differentiable and invertible, we have
$f_Y(y)=f_X(g^{-1}(y))\frac{1}{|g'(g^{-1}(y))|}$
$Z=X+Y$: $f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx$
$=\sum_x P(X=x)P(Y=z-x)$

Variance & expectation

$E[X]=\int x f_X(x) = \sum_x x P(X=x)$
$Var(X)=E[(X-E[X])^2]$
$E[X|A]=\int x f_{X|A}(x)dx$
$E[g(X)|A]=\int g(x)f_{X|A}(x)dx$
$Var(X+Y)=var(X)+var(Y)+2cov(X,Y)$
$Y=\sum_j X_j$:
$E[Y]=\sum_i E[X_i]$
$var(Y)=\sum_i Var(X_i)+2\sum_{i<j}cov(X_i,X_j)$ 0 if indep.
for $X_i \sim Bernoulli(p_i)$,
$E[Y^2]=\sum_i E[X_i]+2\sum_{i<j}E[X_iX_j]$
if X_i indep
$=\sum p_i+2\sum_{i<j}p_ip_j$
if iid,
$=np+n(n-1)p_i^2$

Moments

$M_X(s)=E[e^{sX}]=1+sE[X]+\frac{s^2}{2!}E[X^2]+\dots$
$\left(\frac{d^n}{ds^n}E[e^{sX}]\right)(0)=E[X^n]$
$M_X(0)=1$
$G_X(z)=M_X(\ln z)$, $M_X(s)=G_X(e^s)$
$\frac{\left(\frac{d^n}{ds^n}G_X(z)\right)(0)}{n!}=P(X=n)$
When X is nonnegative, discrete
$\lim_{s\to-\infty}(M_X(s))=P(X=0)$
$Y=aX+B \Rightarrow M_Y(s)=e^{sb}M_X(as)$
$Z=\sum_i X_i \Rightarrow M_Z(s)=\prod_i M_{X_i}(s)$
indep.
$M_Z(s)=\sum_{n=0}^{\infty}(M_X(s))^nP(N=n)$

Known MGFS

Distribution	MGF
$Bernoulli(p)$	$(1-p)+pe^s$
$Bin(n,p)$	$(1-p+pe^s)^n$
$Geo(p)$	$\frac{pe^s}{1-(1-p)e^s}$
$Poisson(p)$	$e^{\lambda(e^s-1)}$
$exp(\lambda)$	$\frac{\lambda}{\lambda-s}$ $\lambda>s$
$uni([a,b])$	$\frac{e^{sb}-e^{sa}}{s(b-a)}$
$N(\mu,\sigma^2)$	$e^{\mu s+\frac{\sigma^2s^2}{2}}$

Covariance: $cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$

$cov(X,X)=var(X)$
$cov(aX_1+bX_2, cY_1+dY_2)$ distribute each like FOIL
$=ac\,cov(X_1,Y_1)+ad\,cov(X_1,Y_2)$
$+bc\,cov(X_2,Y_1)+bd\,cov(X_2,Y_2)$
indep $\Rightarrow$ cov $=0$

classic problems

Ballot: A scores n B scores m s.t. $m<n$
$P($A strictly ahead of B always$) = \frac{n-m}{m+n}$

Gamblers ruin: start at i and move $+1$ or -1 with prob p & $1-p$ respectively. $P($reaching n before $0)$ is

$$\frac{1-\left(\frac{p}{1-p}\right)^{n-i}}{1-\left(\frac{p}{1-p}\right)^{n}}$$

secretary: reject first m candidates, take first better than all prev. optimal $m = ne^{-1}$

Coupon collector:

$$E[T] = n\sum_{i=1}^{n}\frac{1}{i}$$

T trials needed to collect all coupons

Derangement:

$$P(X=n) = n!\sum_{k=0}^{n}\frac{-1^{k}}{k!}$$

neg geometric: B bad G good, $N=G+B$ number of draws til rth good item?

$$P(X=x) = \frac{\binom{x-1}{r-1}\binom{N-x}{G-r}}{C^{N}_{G}}$$

$$E[X] = \frac{N+1}{G+1}$$

$$var(X) = \frac{BG(N+1)}{(G+1)^2(G+2)}$$

Expectation cont.

$$E[X] = \sum_{k=1}^{\infty}P(X\geq k) = \int_0^{\infty}P(X\geq s)ds$$

useful series

$$\sum_{k=0}^{\infty}\frac{x^k}{k!} = e^x$$

$$\sum_{k=0}^{\infty}r^k = \frac{1}{1-r} \quad |r|<1$$

$$\sum_{k=0}^{n}r^k = \frac{1-r^{n+1}}{1-r}, \quad r\neq 1$$

Triangle 1. from HW:

(i) $2\int_0^{1/2}x\,dx = 1/4$

(ii) $\int_0^{1/2}\frac{x}{1-x}dx = \ln 2 - 1/2$

(iii) $2(\ln 2 - 1/2)$

for $S = \sum_{0}^{N}X_i$ $\quad X_i \overset{iid}{\sim} exp(\lambda)$

$N \sim geo(p)$

$S \sim exp(\lambda p)$

for gaussian confidence interval:

$$P\left(\left|\frac{\bar{M}-M}{\sigma/\sqrt{n}}\right|\leq x\right) = \Phi(x)-\Phi(-x)$$

- p confidence interval solve for p

for num trials needed,

$\bar{M} \pm \frac{x}{\sqrt{n}}$ ← x is got by solving confidence interval

$\frac{x}{\sqrt{n}} \leq$ num samples

correlation

$$\rho(x,y) = \frac{cov(x,y)}{\sigma_x\sigma_y}$$

Finding PDF min/max

if $X\sim geo(p)$ $Y\sim geo(q)$

$min(X,Y)\sim geo(p+q-pq)$

if $X_i \overset{iid}{\sim} geo(p)$

$min(X_i) \sim geo(1-(1-p)^n)$

if $X_i \sim exp(\lambda_i)$

$min(X_i) \sim exp(\sum\lambda_i)$

if not geo or poisson

min: $P(X,Y>t)$

Max: $P(X,Y<t)$

Extra dists.

Erlang $\sim(k,\lambda)$ sum of k i.i.d. $exp(\lambda)$

$E[X] = \frac{k}{\lambda}$ $\quad var(X) = \frac{k}{\lambda^2}$

$f(x) = \frac{\lambda^k x^{k-1}e^{-\lambda x}}{(k-1)!}$ $\quad F(x) = 1-\sum_{0}^{k-1}\frac{1}{n!}e^{-\lambda x}(\lambda x)^n$

$X\sim Erlang(k,\lambda)$ $\quad aX\sim Erlang(k,\frac{\lambda}{a})$

Pascal $\sim(k,p)$ sum of k i.i.d. Geo(p)

$E[X] = \frac{k}{p}$ $\quad var(X) = \frac{k(1-p)}{p^2}$

$f(x) = \binom{x-1}{k-1}p^k(1-p)^{x-k}$ $\quad x\in[k,k+1,\dots]$, $k\in\mathbb{N}$

$F(x) = P(X\leq x) = \sum_{n=k}^{x}\binom{n-1}{k-1}p^k(1-p)^{n-k}$

convergence

★ try using log, SLLN or Chebyshevs.

(1) $\lim_{n\to\infty}P(|X_n-X|>\varepsilon)=0$ $\quad X_n \xrightarrow{i.p.} X$

(2) $P(\lim_{n\to\infty}X_n = X) = 1$

Borel cantelli: $\sum P(A_n)<\infty$: A_n occurs finite many times $\sum P(A_n)=\infty$: A_n occurs infinitely often almost surely

Concentration ineq.

Markov: $P(X \geq c) \leq \frac{E[X]}{c}$ $c>0$
for non-neg RV X w/ finite mean

Chebyshevs
$P(|X-\mu| \geq c) \leq \frac{var(X)}{c^2}$
$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

Markov chains

- to show is Markov Chain, show memorylessness
- states:
  transient $P(\text{return})<1$ ex: chains w/ absorbing state (state itself pos. recurrent & states it can reach transient.
  null recurrent: $P(\text{return})=1$ and $E[\text{return time}]=\infty$ ex: symmetric random walk
- Pos recurrent: $P(\text{return})=1$ and $E[\text{return time}]<\infty$. if a chain is finite & irreducible it is pos recurrent
- If a chain is irreducible we have $E[\text{return time to } i \text{ from } i] = \frac{1}{\pi_i}$
- For random walk on undirected graph where you move uniformly to neighbor $\pi_i = \frac{deg(i)}{\sum_j deg(j)}$
- Can you use detailed balance? if yes, try $\frac{\pi_j}{\pi_i} = \frac{P_{ij}}{P_{ji}}$

Poisson processes

- Both forward AND backward interarrival times are exp

Problem solving: in competing poisson process problems, you can simplify into 1 type of event and then look at prob that the event is classified as either type $\frac{\lambda}{\lambda+\mu}$

- Given n events happened in some interval, they are uniformly distributed. If ordered, they will evenly distribute

Statistical inference

MLE estimation:
we find $\max_\theta L(\theta)$ which is $\max_\theta p(x_1,\ldots x_n \mid \theta)$. Maybe take log to simplify. Then differentiate to find max

MAP estimation:
want $P(\theta \mid X)$. Assume you have prior $P(x)$. Then $P(\theta \mid X) \propto P(X \mid \theta) p(\theta)$. Plug in, differentiate & set to 0 to find max

Decision functions:
(1) setup: write out H_0 & H_1
(2) find $\Lambda(x) = \frac{f_1(x)}{f_0(x)}$. If this is inc. in X, we will have $\delta(x) = \begin{cases} 1 & x>c \text{ rej. } H_0 \\ 0 & x \leq c \text{ accept } H_0 \end{cases}$
(3) Usually we will have some bound on $PFA = P(\hat{X}=1 \mid X=0) = \alpha$. write this in terms of decision boundary.
(4) if necessary randomize at boundary. We need to hit α exactly. Write equation of the form $P(\text{rej.}) = P(X \geq b+1) + rP(X=b)$ where we will rej. at boundary w/ prob r. note this inequality may be flipped if Λ dec in x

Statistical inference cont.

If $\Lambda(x)$ is not monotonic then rej H_0 on sets where $\Lambda(x)$ is largest. Randomize on last point if necessary.

MMSE problems: we want to find X from Y by min $E[(X-\hat{X}(Y))^2]$. This is $\hat{X}_{MMSE}(Y) = E[X \mid Y]$

LLSE problems: $\hat{X}(Y) = a+bY$ with same goal as before but $\hat{X}(Y) = a+bY$. Then use $b = \frac{cov(X,Y)}{var(Y)}$ $a = E[X]-bE[Y]$

HW problems:

(1) $X_n \sim \text{Bernoulli}(\frac{1}{n})$

(a) 0 in prob

$P(|X_n - 0| > \varepsilon) = P(X_n = 1) = \frac{1}{n} \to 0$

(b) we have LHS IFF $\exists n \in \mathbb{N}$ where past it all $X_{n+k} = 0$

(c) $P(\lim_{n\to\infty} X_n = 0) \leq \sum_{N=1}^{\infty} P(X_n = 0 \quad n \geq N)$

$= \sum P\left(\bigcap_{n=N}^{\infty} \{X_n = 0\}\right)$

$= \sum \prod P(X_n = 0)$

$= \sum \frac{N-1}{N} \frac{N}{N+1} \frac{N+1}{N+2}$

$= 0$

(2) Convergence in L^p

$P(|X_n - X| \geq \varepsilon) = P(|X_n - X|^p \geq \varepsilon^p)$

$\leq \dfrac{E[|X_n - X|^p]}{\varepsilon^p}$

$X_n \to X \Rightarrow P(|X_n - X| \geq \varepsilon) \to 0$

(3) Conf. intervals Cheb. vs. Chern. vs. CLT

(a) $\mu \in \left(\bar{X}_n - \frac{\sigma}{\sqrt{n}}\frac{1}{\sqrt{\alpha}}, \bar{X}_n + \frac{\sigma}{\sqrt{n}}\frac{1}{\sqrt{\alpha}}\right)$

$= P\left(|\bar{X}_n - \mu| \leq \frac{\sigma}{\sqrt{n}}\frac{1}{\sqrt{\alpha}}\right) = 1 - P(>)$

$\geq 1 - \dfrac{\text{var}(\bar{X}_n)}{\sigma^2/_n (1/\alpha)}$

(b) $P\left(\mu \in \bar{X}_n \pm \frac{\frac{1}{2}}{\sqrt{n}}\sqrt{2\ln\frac{2}{\alpha}}\right) = \begin{array}{l} = 1-\alpha \\ P(| \ | \leq) \end{array}$

$= 1 - P(>)$

$\geq 1 - 2e^{-\ln 2/\alpha}$

$= 1-\alpha$

(c) $P(Z \geq \varepsilon) = P(e^{tZ} \geq e^{t\varepsilon}) = \dfrac{E[e^{tZ}]}{e^{t\varepsilon}}$

$= e^{\frac{1}{2}t^2 - t\varepsilon}$

then optimize over t

$P(Z \geq \varepsilon) \leq e^{-\varepsilon^2/2}$ then union bound

(d) set $\varepsilon = \sqrt{2\ln\frac{2}{\alpha}}$ from prev part

$P\left(|X_n - \mu| \geq \frac{\sigma}{\sqrt{n}}\sqrt{2\ln\frac{2}{\alpha}}\right) \leq \alpha$

then expand, $\mu - X_n$ in middle term

then put back into abs val & done

(4) CLT no upgrade

(a) $\varepsilon > 0$ by union bound

$P(|aX + Y - aX_n - Y_n| > \varepsilon) \leq P(|a(X - X_n)| > \varepsilon/2)$

or

$|Y - Y_n| > \varepsilon/2$

$\leq P(|X - X_n| > \varepsilon/_{2|a|})$

$+ P(|Y - Y_n| > \varepsilon/2)$

(b) $\sqrt{2} Z_{2n} - Z_n = \dfrac{X_{n+1} + \ldots + X_{2n}}{\sqrt{n}} \sim Z_n$ but then

thus $(\sqrt{2} - 1) Z = Z$

(6) CLT $\Rightarrow$ WLLN

(a) $\forall \varepsilon > 0$

$\lim_{n\to\infty} F_{X_n}(c - \varepsilon) = F_c(c - \varepsilon) = 0$

$\lim_{n\to\infty} F_{X_n}(c + \frac{\varepsilon}{2}) = F_c(c + \frac{\varepsilon}{2}) = 1$

$\lim_{n\to\infty} P(|X_n - c| \geq \varepsilon) = \lim_{n\to\infty} P(X_n \leq c - \varepsilon)$

$+ \lim_{n\to\infty} P(X_n \geq c + \varepsilon)$

$\leq \quad >$

$= \lim_{n\to\infty} F_{X_n}(c - \varepsilon) + 1 - F_{X_n}(c + \frac{\varepsilon}{2})$

$= 0$

(b) $Z_n := \frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum X_i - \mu\right) \xrightarrow{\text{in dist}} N(0,1)$

$a_n := \frac{\sigma}{\sqrt{n}} \to 0$ then $Y_n := a_n Z_n$

$= \frac{1}{n}\sum_{i=1}^{n} X_i - \mu \to 0 \quad c = 0 \Rightarrow Y_n$

also $\to 0$ in prob.

(7) coupon collector convergence

$T_n = \sum X_i \quad X_i \sim \text{geo}\left(\frac{n-i+1}{n}\right)$

$E[T_n] = \sum \frac{n}{n-i+1} = n\sum \frac{1}{i} = nH_n$

$\text{var}(T_n) = \sum \text{var}(X_i) \leq \sum\left(\frac{n}{n-i+1}\right)^2$

$= \sum\left(\frac{n}{i}\right)^2 \leq n^2 \sum \frac{1}{i^2}$

(8) ant: $\beta(\cdot) = \dfrac{1 - \rho^n}{\rho^n - \rho^{n+1}}$

(9) when Markov given n?