Introduction

Suppose X is a cts r.v. with pdf f(x) = $\begin{cases} cx^2, -1 < x < 2 \\ 0, otherwise \end{cases}$	$\int_{-1}^{2} cx^2\,dx = 1$. c = 1/3. P(X > 0) = $\int_0^2 x^2/3\,dx = 8/9$
If X is a cts r.v. with pdf f(x) = $\begin{cases} 2e^{-2x}, x > 0 \\ 0, otherwise \end{cases}$ Find median of X. Find pdf of 5X + 10	F(y) = P(X ≤ y) = $\int_0^y 2e^{-2x}\,dx = -e^{-2x}\Big\|_0^y = 1 - e^{-2y}$, y ≥ 0. F(m) = 0.5. $1 - e^{-2m} = 0.5$. m = (ln 2)/2 Let W = 5X + 10. P(W ≤ w) = P(5X + 10 ≤ w) = P(X ≤ (w-10)/5), w ≥ 10 = $1 - e^{-2(w-10)/5}$ Pdf of W. $f_w(w) = \frac{d}{dx}P(W \le w) = \frac{2}{5}e^{-2w/5+4}$, w ≥ 10
X is a cts r.v. with f(x) = $\begin{cases} c, 0 < x < 1 \\ 0, otherwise \end{cases}$	$\int_0^1 c\,dx = 1$. c = 1. P(1/3 < X < 1/2) = $\int_{1/3}^{1/2} 1\,dx = 1/6$
If X ~ cdf F(x), pdf of F(X)?	Let Y = F(X). $F_Y(y) = P(Y \le y) = P(F(X) \le y) = P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$
Generate r.v. from pdf f(x) = $\begin{cases} 2e^{-2x}, x > 0 \\ 0, otherwise \end{cases}$	$F(x) = 1 - e^{-2x}$. $F(x) = 1 - e^{-2X} = U$ ~ uniform(0, 1). X = -ln(1-U)/2 1. Generate U ~ uniform(0,1). 2. Deliver X = -ln(1-U)/2

Expectation and Variance

Find E(X) if X ~ pdf f(x) = $\begin{cases} x^2/3, -1 < x < 2 \\ 0, otherwise \end{cases}$	E(X) = $\int_{-1}^{2} x\frac{x^2}{3}dx = \frac{15}{12}$. E(X²) = $\int_{-1}^{2} x^2\frac{x^2}{3}dx = \frac{33}{15}$. Var(X) = E(X²) - [E(X)]² = $\frac{51}{80}$
If pdf of X, f(x) = $\begin{cases} 1, 0 \le x \le 1 \\ 0, otherwise \end{cases}$, find E(X²) and E(-X)	E(X²) = $\int_0^1 x^2 * 1\,dx = 1/3$. E(-X) = $\int_0^1 (-x) * 1\,dx = -1/2$
Lemma. If Y ≥ 0, E(Y) = $\int_0^\infty P(Y > y)dy$	Proof. $\int_0^\infty P(Y > y)dy = \int_0^\infty \int_y^\infty f_y(x)dx\,dy = \int_0^\infty \int_0^x f_y(x)dy\,dx = \int_0^\infty f_y(x)\int_0^x 1\,dy\,dx = \int_0^\infty f_y(x) * x\,dx$ = E(Y)
If X ~ pdf f(x), then for any real-valued fn g, E[g(X)] = $\int_{-\infty}^{\infty} g(x)f(x)dx$	Proof (for g(x) ≥ 0). E[g(X)] = $\int_0^\infty P(g(X) > y)\,dy = \int_0^\infty \int_{x:g(x)>y} f(x)\,dx\,dy$ $= \int_{x:g(x)>0}\int_0^{g(x)} f(x)\,dy\,dx = \int_{x:g(x)>0} f(x)\int_0^{g(x)} 1\,dy\,dx = \int_{x:g(x)>0} f(x)g(x)dx = \int_{-\infty}^{\infty} f(x)g(x)dx$
E(aX+b) = aE(X) + b	Proof. E(aX+B) = $\int_{-\infty}^{\infty}(ax+b)f(x)dx = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$ = aE(X) + b

Uniform r.v.

If X ~ Uniform(0,1) what is pdf of (b-a)X + a? where a, b are constants and b > a Find E((b-a)X + a)	Let Y = (b-a)X + a. $F_Y(y)$ = P(Y ≤ y) = P((b-a)X + a ≤ y) = P(X ≤ $\frac{y-a}{b-a}$) = $\int_0^{\frac{y-a}{b-a}} 1\,dx = \frac{y-a}{b-a}$, a < y < b. $f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} \frac{1}{b-a}, a < y < b \\ 0, otherwise \end{cases}$. E((b-a)X + a) = (b-a)E(X) + a = (b-a)(1/2) + a = (a+b)/2 OR E(Y) = $\int_a^b y f_Y(y)\,dy$
X ~ Uniform(-1,1) = f(x) = $\begin{cases} 1/2, -1 < x < 1 \\ 0, otherwise \end{cases}$	P(|X| < 1/3) = P(-1/3 < x < 1/3) = $\int_{-1/3}^{1/3} 1/2\,dx = 1/3$ P(X² < 1/3) = P(-1/√3 < x < 1/√3) = $\int_{-1/\sqrt{3}}^{1/\sqrt{3}} 1/2\,dx = 1/\sqrt{3}$

Normal r.v.

Blood cholesterol level is approximately normally distributed with mean 220mg/dL and s.d. 15mg/dL. P(blood cholesterol level < 200)?	Let X = blood cholesterol level. X ~ N(220, 15²) P(X < 200) = P($\frac{X-220}{15} < \frac{200-220}{15}$) = P(Z < -1.33) = 0.0918
If Y ~ Binomial(n = 1000, p = 0.3679), find P(Y ≥ 400) μ = 1000*0.3679 = 367.9. σ^2 = 1000(0.3679)(1-0.3679) = 232.5496. σ = 15.2496	Exact: P(Y ≥ 400) = $\sum_{j=400}^{1000}\binom{1000}{j}0.3679^j(1-0.3679)^{1000-j}$ = 0.01954 Normal approximation: Y ~ N(367.9, 15.2496²). P(Y ≥ 400) = P($\frac{X-367.9}{15.2496} \ge \frac{400-367.9}{15.2496}$) = P(Z ≥ 2.07) = 1 - 0.9808 = 0.0192
If X ~ N (μ, σ^2), then $\frac{X-\mu}{\sigma}$ ~ N(0,1)	Proof. Let Y = $\frac{X-\mu}{\sigma}$. $F_Y(y)$ = P(Y ≤ y) = P($\frac{X-\mu}{\sigma}$ ≤ y) = P(X ≤ $\mu + \sigma y$) = $F_X(\mu + \sigma y)$ $f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\mu+\sigma y) = f_X(\mu+\sigma y)\frac{d}{dy}(\mu+\sigma y) = f_X(\mu+\sigma y) * \sigma = \frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\left(\frac{\mu+\sigma y-\mu}{\sigma}\right)^2} * \sigma = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$, -∞ < y < ∞. Y ~ N(0, 1)

Exponential r.v.

Lifetime of a light bulb is exponential r.v. with mean 3 years. Let X = lifetime of light bulb. E(X) = 3. $\lambda = \frac{1}{3}$. f(x) = $\frac{1}{3}e^{-x/3}$, x > 0.	P(X < 3) = $\int_0^3 \frac{1}{3}e^{-x/3}\,dx = 0.63$
Show exponential r.v. is memoryless. X ~ Exp(λ). f(x) = $\lambda e^{-\lambda x}$, x > 0	P(X > s) = $\int_s^\infty \lambda e^{-\lambda x}\,dx = e^{-\lambda s}$. P(X > s+t) = $e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t}$ = P(X > s)P(X > t)
Post office is staffed by 2 clerks. Suppose that when C enters office, he sees A being served by 1 clerk, B by the other. C service will begin when either A or B leaves. If amt of time clerk spend with customer is exponentially dist with parameter λ, what is the prob that C is last to leave?	By memoryless property, time clerk spend with C = time clerk spend with A or B (depending on who haven't finish). So prob = 1/2.
If X ~ Exp(λ), then E(X) = $\frac{1}{\lambda}$ and Var(X) = $\frac{1}{\lambda^2}$	Proof. E(Xⁿ) = $\int_0^\infty x^n\lambda e^{-\lambda x}\,dx = x^n(-e^{-\lambda x})\Big\|_0^\infty + n\int_0^\infty x^{n-1}e^{-\lambda x}\,dx$ (by parts) = $\frac{n}{\lambda}\int_0^\infty x^{n-1}\lambda e^{-\lambda x}\,dx = \frac{n}{\lambda}E(X^{n-1})$ E(X) = = $\frac{1}{\lambda}$. E(X²) = = $\frac{2}{\lambda}$E(X) = $\frac{2}{\lambda^2}$. Var(X) = $\frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$

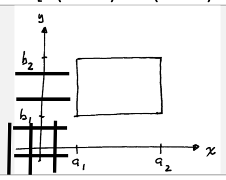
Other cts dist

Γ(α) = (α - 1)Γ(α - 1)!, α > 1	Proof. Γ(α) = $\int_0^\infty e^{-y}y^{\alpha-1}\,dy = -e^{-y}y^{\alpha-1}\Big\|_0^\infty + \int_0^\infty e^{-y}(\alpha-1)y^{\alpha-2}\,dy$ (by parts) = (α-1) $\int_0^\infty e^{-y}y^{\alpha-2}\,dy$ = (α-1)Γ(α - 1)
If α is an int, say α = n, then Γ(n) = (n-1)!. Note Γ(1) = $\int_0^\infty e^{-y}\,dy = 1$	Γ(n) = (n-1)Γ(n-1) = ... = (n-1)(n-2)...(3)(2)(1)Γ(1) = (n-1)!
X_i ~ Exp(λ), i = 1,2,... and X_i are indep. Then $T_n = X_1 + X_2 + ... + X_n$ ~ Gamma(n, λ) Let T_n = time which nth event occurs, N(t) = num of events in time period [0,t]. Note {T_n ≤ t} = {N(t) ≥ n}	P(T_n ≤ t) = P(N(t) ≥ n) = $\sum_{j=n}^{\infty} P(N(t) = j)$. N(t) ~ Poisson($\lambda$t) P($T_n$ ≤ t) = $\sum_{j=n}^{\infty}\frac{e^{-\lambda t}(\lambda t)^j}{j!}$. f(t) = $\frac{d}{dt}$P(T_n ≤ t) = $\sum_{j=n}^{\infty}\frac{e^{-\lambda t}(j)(\lambda t)^{j-1}(\lambda) + (-\lambda)e^{-\lambda t}(\lambda t)^j}{j!} = \sum_{j=n}^{\infty}\frac{\lambda e^{-\lambda t}(\lambda t)^{j-1}}{(j-1)!} -$ $\sum_{j=n}^{\infty}\frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!} = \left(\frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!} + \frac{\lambda e^{-\lambda t}(\lambda t)^n}{n!} + \ldots\right) - \left(\frac{\lambda e^{-\lambda t}(\lambda t)^n}{n!} + \frac{\lambda e^{-\lambda t}(\lambda t)^{n+1}}{(n+1)!} + \ldots\right) = \frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}$
X ~ Cauchy(θ). E(Xⁿ) DNE for n = 1,2,...	Consider θ = 0. E(X) = $\int_{-\infty}^{\infty} xf(x)\,dx = \int_{-\infty}^{\infty} x\frac{1}{\pi}\frac{1}{1+x^2}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2x}{1+x^2}\,dx = \frac{1}{2\pi}\ln(1+x^2)\Big\|_{-\infty}^{\infty}$ = ∞ - ∞ = undefined

Extra

X is a cts r.v. w pdf f(x) = $\frac{x}{8} + \frac{1}{8}$ for -1 < x < 3.	P(X > 0) = $\int_0^3 \frac{x}{8} + \frac{1}{8}\,dx = \frac{15}{16}$. P(X² < 1) = P(-1 < X < 1) = $\int_{-1}^{1}\frac{x}{8} + \frac{1}{8}\,dx = \frac{1}{4}$. E(X) = $\int_{-1}^{3}\frac{x^2}{8} + \frac{x}{8}\,dx = \frac{5}{3}$
A point is chosen at random on a line of length L. Interpret this statement and find prob ratio of shorter to longer segment is < 1/4	Let X denote point chosen. X ~ Uniform(0, L). P(min($\frac{X}{L-X}, \frac{L-X}{X}$) < $\frac{1}{4}$) = 1 - P(min($\frac{X}{L-X}, \frac{L-X}{X}$) > $\frac{1}{4}$) = 1 - P($\frac{X}{L-X} > \frac{1}{4}, \frac{L-X}{X} > \frac{1}{4}$) = 1 - P(X > $\frac{L}{5}$, X < $\frac{4L}{5}$) = 1 - P($\frac{L}{5}$ < X < $\frac{4L}{5}$) = 1 - 3/5 = 2/5
Arrived at bus stop at 10am. Bus arrive at some time uniformly dist btw 10am and 10.30 am. Let X = arrival of bus, X ~ U(0,30)	P(X ≥ 10) = $\int_{10}^{30} 1/30\,dx = 2/3$. P(X ≥ 25 | X ≥ 15) = $\frac{P(X \ge 25, X \ge 15)}{P(X \ge 15)} = \frac{P(X \ge 25)}{P(X \ge 15)} = \frac{5/30}{15/30}$ = 1/3
Fire station is located along road of length a. If fire occur at points uniformly chosen on (0,a), where should station be to minimise expected dist from fire? i.e. choose t s.t. E(|X-t|) is minimized.	Let s(t) = E|X-t| = $\int_0^a \|x-t\|\frac{1}{a}\,dx = \frac{1}{a}\int_0^t t - x\,dx + \frac{1}{a}\int_t^a x - t\,dx = \frac{t^2+(a-t)^2}{2a}$. s'(t) = $\frac{2t-2(a-t)}{2a}$. Let s'(t) = 0, then t = a/2. Checking, s''(a/2) > 0, so t = a/2 gives min value

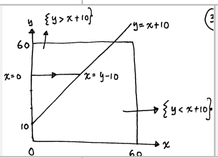
Let X be a r.v..Let a be a real num, show $E(X-a)^2 = var(X) + [a-E(X)]^2$. What is min value of $E(X-a)^2$?	$E(X-a)^2 = E((X-\mu) + (\mu - a))^2 = E((X-\mu)^2 + 2(X-\mu)(\mu - a) + (\mu - a)^2) = E(X-\mu)^2 + 2(\mu - a)E(X-\mu) + (\mu - a)^2 = var(X) + (a-E(X))^2$. Min value of $E(X-a)^2 = var(X)$ and occurs at $a = E(X)$
Let Z be standard normal r.v.. For x > 0, show $P(Z > x) = P(Z < -x)$, $P(\lvert Z\rvert > x) = 2P(Z > x)$, $P(\lvert Z\rvert < x) = 2P(Z < x) - 1$	Note -Z also standard normal r.v.. $P(Z > x) = P(-Z < -x) = P(Z < -x)$. $P(\lvert Z\rvert > x) = P(Z > x) + P(Z < -x) = 2P(Z > x)$. $P(\lvert Z\rvert < x) = P(-x < Z < x) = P(Z < x) - P(Z \le -x) = P(Z < x) - P(Z \ge x) = P(Z < x) - [1 - P(Z < x)] = 2P(Z < x) - 1$
If X is normal r.v. w $\mu = 10$ and $\sigma^2 = 36$, compute P(X > 5), P(4 < X < 16)	$Y = (X-10)/6 \sim N(0,1)$. $P(X \ge 5) = P(Y \ge -5/6) = P(Y < 5/6) = 0.7967$. P (\$ < X < 16) = $P(-1 < Y < 1) = 2P(Y < 1) - 1 = 0.6826$.
Annual rainfall is normally distributed with $\mu = 40$ and $\sigma^2 = 16$. Prob that it will take over 10 years before a year occurs with rainfall > 50 inches. Assumptions made?	In a year, $P(X > 50) = P(\frac{X-40}{4} > \frac{50-40}{4}) = P(Z > 2.5) = 1-0.99379$ Assume events of observing rainfall greater than 50 inches in each year is indep. Then waiting time T until year w rainfall > 50 inches is geometric r.v. w p = 1 - 0.99379. $P(T > 10) = (1-p)^{10}$
1000 indep rolls of a fair die is made. Compute approximation to prob that num 6 will appear btw 150 and 200 times inclusively. If num 6 appears exactly 200 times, find prob that num 5 appear less than 150 times	Total num of sixes rolled, $X \sim Binomial(1000,1/6)$. Using normal approximation, $X \sim N(np, npq) = (1000/6, 5000/36)$. $P(150 \le X \le 200) = P(149.5 < X < 200.5) \approx P(-1.46 < Z < 2.87) = 0.9258$ If 6 appear 200 times, prob 5 appear on other 800 rolls is 1/5. $Y \sim Binomial(800, 1/5) \approx N(800/5, 3200/25)$. $P(Y < 150) = P(Y < 149.5) \approx P(Z < (149.5-160)/sqrt(128)) = P(Z < - 0.93) = 0.1762$
In 10,000 indep toss, coin lands heads 5800 times. Is coin fair?	Suppose coin is fair, then $X \sim Binomial(10000,1/2)$. Using normal approximation, $P(X \ge 5800) = P(X \ge 5799.5) \approx P(Z \ge (5799.5-5000)/sqrt(2500)) = P(Z \ge 15.99) \approx 0$. So unlikely that coin is fair.
Time required to repair machine is an exponentially dist r.v. w $\lambda = 1/2$. Prob that repair > 2 hours. Conditional prob repair takes 10 hours, given its duration exceeds 9 hours.	Let T denote repair time. $P(T > 2) = \int_2^\infty \frac{1}{2}e^{-t/2}\ dt = e^{-1}$ $P(T > 10 \mid T > 9) = P(T > 1)$ (by memoryless property) $= e^{-1/2}$
Y is exponentially dist r.v. w $\lambda = 1$. Compute pdf of r.v. X = logY	$F_X(x) = P(X \le x) = P(\log Y \le x) = P(Y \le e^x) = F_Y(e^x)$. $f_X(x) = f_y(e^x)e^x = e^{-e^x}e^x = e^{-(e^x-x)}$
Weibull dist. Let $\alpha, \beta > 0$ and $v \in \mathbb{R}$. Suppose X is exponentially dist w mean 1. Find pdf and dist fn of Y where $Y = \alpha X^{1/\beta} + v$	Note that Y takes in value from (v, ∞). Let y > v, then $F_Y(y) = P(Y \le y) = P(\alpha X^{1/\beta} + v \le y) = P(X \le \left(\frac{y-v}{\alpha}\right)^\beta) = F_X(\left(\frac{y-v}{\alpha}\right)^\beta) = 1 - e^{-\left(\frac{y-v}{\alpha}\right)^\beta}$. So, $f_Y(y) = f_x(\left(\frac{y-v}{\alpha}\right)^\beta)\ \beta \left(\frac{y-v}{\alpha}\right)^{\beta-1}\frac{1}{\alpha}$. $f_Y(y) = \begin{cases}\frac{\beta}{\alpha}\left(\frac{y-v}{\alpha}\right)^{\beta-1} f_X\left(\left(\frac{y-v}{\alpha}\right)^\beta\right), y > v \\ 0, otherwise\end{cases}$
Let $Y = \left(\frac{X-v}{\alpha}\right)^\beta$. Show if X is a Weibull r.v. w params v, α, β, then Y is an exponential r.v. w $\lambda = 1$.	$P(Y \le y) = P(\left(\frac{X-v}{\alpha}\right)^\beta \le y) = P(X \le v + \alpha y^{1/\beta}) = 1 - e^{-\left(\frac{v + \alpha y^{1/\beta} - v}{\alpha}\right)^\beta} = 1 - e^{-y}$ (cdf of exp w $\lambda = 1$)
Trains headed for A arrive at 15-mins intervals starting from 7am, while trains heading for B arrive at 15-mins intervals starting from 7.05am. If passenger arrive at station at time uniformly dist btw 7 and 8am, and get on 1st train that arrives, prob go to A? If arrive from 7.10 to 8.10am?	P(A) = P(5 < X < 15 or 20 < X < 30 or 35 < X < 45 or 50 < X < 60) = 40/60 = 2/3 since $X \sim Uniform(0,60)$ P(A for 7.10 to 8.10) = P(10 < X < 15 or 20 < X < 30 or 35 < X < 45 or 50 < X < 60 or 65 < X < 70) = 2/3
P(success) = .95. Approximate prob at most 10 of next 150 items produces are unacceptable. Let X denote num of unacceptable items among next 150 produced.	$X \sim Binomial(150, 0.05) \approx N(150*0.5 = 7.5, 150*.5*.95 = 7.125)$ $P(X \le 10) = P(X \le 10.5)$ (continuity correction) $= P(Z \le \frac{10.5-7.5}{\sqrt{7.125}}) \approx P(Z \le 1.1239) = .8695$
Curr price of stock is s. After 1 period, price is either us w prob p or ds w prob 1-p. Assume successive movements are indep, approximate prob stock price will be up at least 30% after next 1000 periods if u = 1.012, d = 0.990, p = .52.	Let X = num of 1000 time periods in which stock increase. Price at end: $su^Xd^{1000-X} = sd^{1000}(u/d)^X$ We need $sd^{1000}(u/d)^X > 1.3s$ OR $d^{1000}(u/d)^X > 1.3$ OR $X > \frac{log(1.3)-1000log(d)}{log(u/d)} = 469.2$. Thus, we need ≥ 470 periods. Using normal approximation for $X \sim Binomial(1000, .52)$ $P(X \ge 470) = P(X > 469.5) = P(Z > \frac{469.5-1000(.52)}{\sqrt{1000(.52)(.48)}}) \approx P(Z > -3.196) \approx .9993$
Show $E[Y] = \int_0^\infty P(Y > y)dy - \int_0^\infty P(Y < -y)dy$	$\int_0^\infty P(Y < -y)dy = \int_0^\infty \int_{-\infty}^{-y} f_Y(x)\,dx\,dy = \int_{-\infty}^0 \int_0^{-x} f_Y(x)\,dy\,dx = \int_{-\infty}^0 [yf_Y(x)] \Big\rvert_0^{-x} dx = -\int_{-\infty}^0 xf_Y(x)\,dx$ Similarly, $\int_0^\infty P(Y > y)dy = \int_0^\infty xf_Y(x)\,dx$. So, $\int_0^\infty P(Y > y)dy - \int_0^\infty P(Y < -y)dy = \int_0^\infty xf_Y(x)\,dx + \int_{-\infty}^0 xf_Y(x)\,dx = \int_{-\infty}^\infty xf_Y(x)\,dx = E[Y]$
Use the result that for a nonnegative r.v. Y, $E[Y] = \int_0^\infty P(Y > t)dt$ to show for a nonnegative r.v. X, $E[X^n] = \int_0^\infty nx^{n-1}P(X > x)dx$.	Let $t = x^n$, then $\frac{dt}{dx} = nx^{n-1}$. $E[X^n] = \int_0^\infty P(X^n > t)dt = \int_0^\infty P(X^n > x^n)nx^{n-1}dx = \int_0^\infty nx^{n-1}P(X > x)dx$
Let X be a r.v. that takes on values btw 0 and c. i.e. $P(0 \le X \le c) = 1$. Show $var(X) \le c^2/4$	Since $0 \le X \le c$, then $X^2 \le cX$, so $E[X^2] \le E[cX]$ $Var(X) = E[X^2] - (E[X])^2 \le E[cX] - (E[X])^2 = cE[X] - (E[X])^2 = E[X](c - E[X]) = c^2[a(1-a)]$ (where $a = E[X]/c$) $\le c^2/4$ (since max of a(1-a) = 1/4 using differentiation)
If X is an exponential r.v. w mean $1/\lambda$, show $E[X^k] = \frac{k!}{\lambda^k}$, k = 1,2,...	$E[X^k] = \int_0^\infty x^k \lambda e^{-\lambda x}\,dx = \lambda^{-k}\int_0^\infty \lambda e^{-\lambda x}(\lambda x)^k\,dx = \frac{\Gamma(k+1)}{\lambda^k}\int_0^\infty \frac{\lambda e^{-\lambda x}(\lambda x)^k}{\Gamma(k+1)}dx = \frac{\Gamma(k+1)}{\lambda^k}(1)$ (gamma pdf) $= \frac{k!}{\lambda^k}$
Show $\Gamma(1/2) = \sqrt{\pi}$. Let $y = \sqrt{2x}$. $\frac{dy}{dx} = \frac{1}{\sqrt{2x}}$	$\Gamma(1/2) = \int_0^\infty e^{-x}x^{1/2-1}\,dx = \int_0^\infty e^{-y^2/2}x^{-1/2}\ \sqrt{2x}\,dx = \sqrt{2}\int_0^\infty e^{-y^2/2}\,dy = 2\sqrt{\pi}\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}\,dy = 2\sqrt{\pi}\,P(Z > 0) = 2\sqrt{\pi}(1/2) = \sqrt{\pi}$
Find pdf of $Y = e^X$ when X is normally dist w params μ and σ^2. r.v. Y is said to have a lognormal dist w params μ and σ^2.	$F_Y(x) = P(Y \le x) = P(e^X \le x) = P(X \le \ln x) = F_X(\ln x)$. $f_Y(x) = f_X(\ln x)(1/x) = \frac{1}{x\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2}$

Joint Distribution Fn	
Marginal cdf of X, $F_X(x) = \lim_{y\to\infty} F(x,y)$	Proof. $F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = P(\lim_{y\to\infty}\{X \le x, Y \le y\}) = \lim_{y\to\infty} P(X \le x, Y \le y) = \lim_{y\to\infty} F(x,y)$
$P(X > a, Y > b) = 1 - F_X(a) - F_y(b) + F(a,b)$	Proof. $P(X > a, Y > b) = 1 - P(\{X > a, Y \ge b\}^C) = 1 - P(\{X > a\}^C \cup \{Y > b\}^C)$ (since $(A\cap B)^C = A^C \cup B^C$) $= 1 - P(\{X \le a\} \cup \{Y \le b\}) = 1 - [P(X \le a) + P(Y \le b) - P(X \le a, Y \le b)]$ (since $P(A\cup B) = P(A) + P(B) - P(AB)$) $= 1 - F_X(a) - F_y(b) + F(a,b)$
$P(a_1 \le X \le a_2, b_1 \le Y \le b_2) = F(a_2, b_2) + F(a_1, b_1) - F(a_1, b_2) - F(a_2, b_1)$	Proof. 
Suppose 2 balls are chosen w/o replacement from urn consisting of 4W, 6B balls. Let $X_i = 1$ if i^{th} ball selected is white and 0 otherwise. Find joint pmf of (X_1, X_2)	$P(X_1 = 0, X_2 = 0) = (6/10)(5/9) = 1/3$... pmf of X_1. $P(X_1 = 0) = 1/3 + 4/15 = 3/5$ $P(X_1 = 1) = 4/15 + 2/15 = 2/5$ Can find pmf of X_2 as well

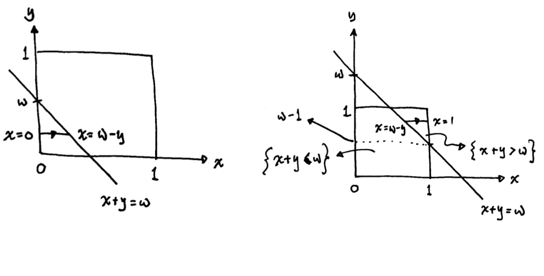
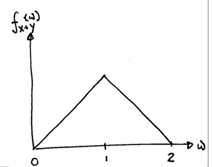
$X_1 \backslash X_2$	0	1
0	1/3	4/15
1	4/15	2/15

$f(x, y) = \begin{cases} 2e^{-x}e^{-2y}, 0 < x < \infty, 0 < y < \infty \\ 0, otherwise \end{cases}$ Find probs. Find pdf of X and pdf of Y	P(X > 1, Y < 2) = $\int_0^2 \int_1^\infty 2e^{-x}e^{-2y}dxdy = \int_0^2 2e^{-2y}[\int_1^\infty e^{-x}dx]dy = \int_0^2 2e^{-2y}[-e^{-x}]\Big\vert_1^\infty dy$ = $\int_0^2 2e^{-2y}[-e^{-1}]dy$ = $-e^{-1}\int_0^2 -2e^{-2y}dy = -e^{-1}[e^{-2y}]\Big\vert_0^2 = -e^{-1}(e^{-4} - 1) = e^{-1}(1-e^{-4})$ P(X < Y) = $\int_0^\infty \int_0^y 2e^{-x}e^{-2y}dxdy$ = ... = 1/3. P(Y > 2) = $\int_0^\infty \int_2^\infty 2e^{-x}e^{-2y}dydx$ = ... = e^{-4} $f_X(x) = \int_0^\infty 2e^{-x}e^{-2y}dy = ... = \begin{cases} e^{-x}, x > 0 \\ 0, otherwise \end{cases}$. $f_Y(y) = \int_0^\infty 2e^{-x}e^{-2y}dx = ... = \begin{cases} 2e^{-2y}, y > 0 \\ 0, otherwise \end{cases}$
Joint pdf of X and Y, $f(x,y) = \begin{cases} e^{-(x+y)}, 0 < x < \infty, 0 < y < \infty \\ 0, otherwise \end{cases}$ Find pdf of X/Y. Let W = X/y	$F_W(w) = P(W \le w) = P(X/Y \le w) = \int_0^\infty \int_0^{wy} e^{-x-y}dxdy = \int_0^\infty e^{-y}\int_0^{wy} e^{-x}dx\,dy = ... = 1 - \frac{1}{w+1}$, w > 0 (since x > 0, y > 0, x/y > 0). $f_W(w) = \frac{d}{dw}F_W(w) = \frac{1}{(w+1)^2}$

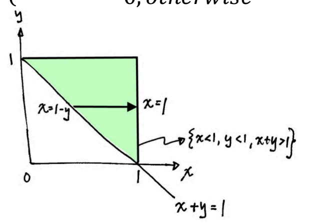
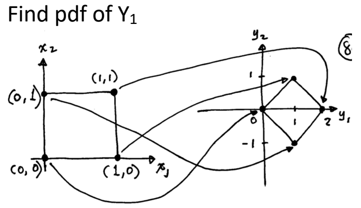
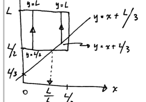
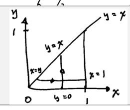
Indep r.v.

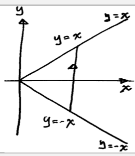
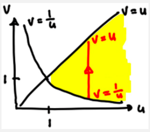
Suppose that n+m indep trials, with P(success) = p. Let X = num of successes in 1st n trials, Y = num of success in last m trials, Z = total success in n+m trials	$P(X = x, Y = y) = \binom{n}{x}p^x(1-p)^{n-x}\binom{m}{y}p^y(1-p)^{m-y} = P(X = x)P(Y = y)$, $0 \le x \le n$, $0 \le y \le m$. So X and Y are indep Z = X + Y. $P(X = x, Z = z) = P(X = x, X + Y = z) = P(X = x, Y = z - x) = \binom{n}{x}p^x(1-p)^{n-x}\binom{m}{z-x}p^{z-x}(1-p)^{m-z+x}$. So $P(X = x, Z = z) \ne P(X = x)P(Z = z) = \binom{n}{x}p^x(1-p)^{n-x}\binom{n+m}{z}p^z(1-p)^{n+m-z}$. So X and Z are dependent
Num of ppl entering post office in a day is Poisson r.v. w paremeter λ. If P(male) = p, P(female) = 1-p. Show num of males and females entering office are indep Poisson r.v. w respective parameters λp and λ(1-p)	Let X = num of males entering, Y = num of females. Given X + Y ~ Poisson(λ). $P(X + Y = i+j) = e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$. $P(X = i, Y = j \mid X + Y = i+j) = \binom{i+j}{i}p^i(1-p)^j$ $P(X = i, Y = j) = P(X = i, Y = j \mid X + Y = i+j)P(X + Y = i+j) + P(X = i, Y = j \mid X + Y \ne i+j)P(X + Y \ne i+j)$ = $\binom{i+j}{i}p^i(1-p)^je^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!} + 0 = \frac{(i+j)!}{i!j!}(\lambda p)^i[\lambda(1-p)]^j\frac{e^{-\lambda}}{(i+j)!} = \frac{e^{-\lambda p}(\lambda p)^i}{i!}\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}$ = P(X=i)P(Y=j) So X and Y indep. $P(X=i) = \frac{e^{-\lambda p}(\lambda p)^i}{i!}$, i = 0,1,2,.... X ~ Poisson(λp). Y ~ Poisson(λ(1-p))
Man and woman meet. If each person independently arrives at a time uniformly distributed btw 12 and 1pm, find prob 1st to arrive has to wait longer than 10 mins.  	Let X = time past 12pm man arrives, Y = time past 12pm woman arrives. X ~ Uniform(0,60), Y ~ Uniform(0,60). $f(x,y) = f_x(x)f_y(y) = (1/60)(1/60) = 1/60^2$, 0 < x < 60, 0 < y < 60 P(Y > X + 10 or X > Y + 10) = P(Y > X + 10) + P(X > Y + 10) = 2P(Y > X + 10) (by symmetry) = $2\int_{10}^{60}\int_0^{y-10}\frac{1}{60^2}dx\,dy$ = 25/36 OR $2\int_0^{50}\int_{x+10}^{60}\frac{1}{60^2}dy\,dx$
Buffon's needle problem. Table has equidistant parallel lines at distance D apart. Needle of length L, where L ≤ D, is randomly thrown on table. Prob that needle will intersect one of the lines?	Needle will intersect if $\frac{L}{2}\cos\theta > X$, where X = dist from middle of needle to nearest // line. Note X ~ Uniform(0, D/2), θ ~ Uniform(0, π/2) and X and θ are indep. $f(x,y) = f_x(x)f_\theta(y) = 1/(D/2) * 1/(\pi/2) = 4/(D\pi)$, $0 \le x \le D/2$, $0 \le y \le \pi/2$ $P(X < \frac{L}{2}\cos\theta) = \int_o^{\pi/2}\int_0^{\frac{L}{2}\cos\theta}\frac{4}{\pi D}dx\,dy = \frac{2L}{\pi D}$. Then $\pi = \frac{2L}{D}\frac{1}{P(X<(L/2)cos\theta}= \frac{2L}{D}\frac{1}{P(needle\ intersect\ a\ line)}$ By throwing needle N times, find num of times needle intersect line, then $\pi = \frac{2L}{D}\frac{N}{n}$
X and Y indep iff their joint pdf/pmf can be expressed as f(x, y) = g(x)h(y), -∞ < x < ∞, -∞ < y < ∞	Proof. ⇒: X and Y indep ⇒ $f(x,y) = f_x(x)f_y(y)$. ⇐: f(x,y) = h(x)g(y). $1 = \int_{-\infty}^\infty\int_{-\infty}^\infty f(x,y)\,dx\,dy$ = $\int_{-\infty}^\infty h(x)dx\int_{-\infty}^\infty g(y)\,dy = c_1c_2$. $f_x(x) = \int_{-\infty}^\infty f(x,y)\,dy = h(x)\int_{-\infty}^\infty g(y)\,dy = c_2h(x)$. Similarly, $f_y(y) = c_1g(y)$. So, $f(x,y) = h(x)g(y) = \frac{f_x(x)}{c_2}\frac{f_Y(y)}{c_1} = f_x(x)f_y(y)$ (since c_1c_2 = 1). So X and Y indep.
$f(x, y) = \begin{cases} 6e^{-2x}e^{-3y}, \begin{matrix} 0 < x < \infty \\ 0 < y < \infty \end{matrix} \\ 0, otherwise \end{cases}$. X and Y indep?	$f(x,y) = 6e^{-2x}e^{-3y}I_x(x)I_y(y) = 6e^{-2x}I(x) * e^{-3y}I(y) = g(x)h(y)$ for $\begin{matrix} -\infty < x < \infty \\ -\infty < y < \infty \end{matrix}$, where $I_x(x) = \begin{cases} 1, 0 < x < \infty \\ 0, otherwise \end{cases}$, same for $I_y(y)$. So X and Y indep since f(x, y) = g(x)h(y), **-∞ < x < ∞, -∞ < y < ∞**
$f(x, y) = \begin{cases} 24xy, \begin{matrix} 0 < x < 1 \\ 0 < y < 1 \end{matrix}, 0 < x + y < 1 \\ 0, otherwise \end{cases}$. Are X and Y indep?	f(x,y) ≠ h(x)g(y) for all x,y -∞ < x < ∞, -∞ < y < ∞. Define I(x,y) = $= \begin{cases} 1, \begin{matrix} 0 < x < 1 \\ 0 < y < 1 \end{matrix}, 0 < x + y < 1 \\ 0, otherwise \end{cases}$ f(x,y) = 24xyI(x,y) -∞ < x < ∞, -∞ < y < ∞ which cannot be factored as h(x)g(y), so not indep.
Let X_1, X_2, ... be seq of indep and identically distributed cts r.v. and suppose we observe these r.v. in seq. If $X_n > X_i$ for each i = 1,2,...,n-1, then we say X_n is a record value. Let A_n = {X_n be record value}. Is A_{n+1} indep of A_n?	E.g. $P(A_6)$ = any one of the 6 x's can be the largest, so $P(A_6)$ = 1/6 $P(A_{n+1} \mid A_n) = P(A_{n+1})$? If it is, then indep. But this is diff to calculate Consider $P(A_n \mid A_{n+1}) = P(A_n)$ (since A_{n+1} is a future event) = 1/n. So A_n is indep of A_{n+1} By symmetric relation of independence, A_{n+1} also indep of A_n

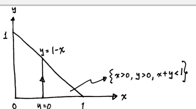
Sum of indep r.v.

X ~ Uniform(0,1), Y ~ Uniform(0,1), X + Y ~ triangular  	$f_x(x) = \begin{cases} 1, 0 < x < 1 \\ 0, otherwise \end{cases}$, $f_Y(y) = \begin{cases} 1, 0 < y < 1 \\ 0, otherwise \end{cases}$. $f(x,y) = f_x(x)f_y(y) = \begin{cases} 1, 0 < x < 1, 0 < y < 1 \\ 0, otherwise \end{cases}$ Let W = X+Y. $F_W(w) = P(W \le w) = P(X + Y \le w)$ For 0 < w < 1, $P(X+Y \le w) = P(X \le w - y) = \int_0^w\int_0^{w-y} f(x,y)\,dx\,dy = \int_0^w\int_0^{w-y} 1\,dx\,dy = w^2/2$ For 1 < w < 2, $P(X + Y \le w) = 1 - P(X + Y > w) = \int_{w-1}^1\int_{w-y}^1 1\,dx\,dy = 2w - w^2/2 - 1$ $F_{X+Y}(w) = \begin{cases} 0, w < 0 \\ \frac{w^2}{2}, 0 < w < 1 \\ 2w - \frac{w^2}{2} - 1, 1 < w < 2 \\ 1, w > 2 \end{cases}$, $f_{X+Y}(w) = \begin{cases} w, 0 < w < 1 \\ 2 - w, 1 < w < 2 \\ 0, otherwise \end{cases}$ 
3. Z_i ~ N(0, 1), i = 1,...,n ⇒ $\sum_{i=1}^n Z_i^2 \sim \chi_n^2$ (chi-square w n deg of freedom) Note pdf of Gamma($\frac{1}{2}, \frac{1}{2}$) = $\frac{\frac{1}{2}e^{-\frac{1}{2}y}\left(\frac{1}{2}y\right)^{\frac{1}{2}-1}}{\Gamma(\frac{1}{2})}, y \ge 0$ $\Gamma(\frac{1}{2}) = \int_0^\infty e^{-y}y^{\frac{1}{2}-1}\,dy = \sqrt{\pi}$	Proof. Z ~ N(0,1). $f_Z(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$, -∞ < z < ∞. Let Y = Z^2. $F_Y(y) = P(Y \le y) = P(Z^2 \le Y) = P(-\sqrt{y} \le z \le \sqrt{y}) = P(Z \le \sqrt{y}) - P(Z \le -\sqrt{y}) = F_Z(\sqrt{y}) - F_Z(-\sqrt{y})$. $f_Y(y) = \frac{d}{dy}F_Y(y) = f_z(\sqrt{y})\frac{1}{2}y^{-\frac{1}{2}} - f_Z(-\sqrt{y})\left(-\frac{1}{2}y^{-\frac{1}{2}}\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\sqrt{y})^2}\frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(-\sqrt{y})^2}\frac{1}{2\sqrt{y}}$ = $\frac{1}{2\sqrt{\pi}}e^{-\frac{1}{2}y}\left(\frac{y}{2}\right)^{\frac{1}{2}-1}$, y > 0. So Z^2 ~ Gamma($\frac{1}{2}, \frac{1}{2}$). Using result 1, $Z_1^2 + Z_2^2 + ... + Z_n^2$ ~ Gamma($\frac{n}{2}, \frac{1}{2}$)
5. X ~ Poisson(λ_1), Y ~ Poisson(λ_2) ⇒ X + Y ~ Poisson($\lambda_1 + \lambda_2$)	Proof. P(X + Y = n) = P(X = 0, Y = n) + P(X = 1, Y = n-1) + ... + P(X = n, Y = 0) = $\sum_{k=0}^n P(X = k, Y = n - k)$ = $\sum_{k=0}^n P(X = k)P(Y = n - k)$ (since indep) = $\sum_{k=0}^n \frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^n\frac{\lambda_1^k\lambda_2^{n-k}}{k!(n-k)!}$ = $\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^n\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k} = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1 + \lambda_2)^n$ (binomial expansion)

Conditional dist for discrete case

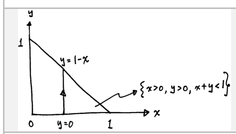
If X and Y are indep Poisson r.v. w parameters λ_1, λ_2. Conditional dist of X given X + Y = n?	X ~ Poisson(λ_1), Y ~ Poisson(λ_2), X + Y ~ Poisson($\lambda_1 + \lambda_2$). P(X = k|X + Y = n) = $\frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} = \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}$ (since indep) = $\frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!} * \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}}{\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n}$ = $\frac{n!}{k!(n-k)!}\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n} = \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} = \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(1-\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}$. X|X+Y=n ~ Binomial(n, $\frac{\lambda_1}{\lambda_1+\lambda_2}$)
Conditional dist for cts case	
Joint pdf of X and Y is f(x,y) = $\begin{cases} c(x+y^2), x<1, y<1, x+y>1 \\ 0, otherwise\end{cases}$ 	$1 = \int_0^1\int_{1-y}^1 c(x+y^2)\,dx\,dy$... c = $\frac{12}{7}$ marginal pdf of y, f_Y(y) = $\int_{1-y}^1 \frac{12}{7}(x+y^2)\,dx = \frac{12}{7}$(y³ + y - y²/2), 0 < y < 1 conditional pdf of X given Y = y, $f_{X\|Y}$(x|y) = $\frac{f(x,y)}{f_Y(y)} = \frac{\frac{12}{7}(x+y^2)}{\frac{12}{7}(y^3+y-y^2/2)} = \frac{(x+y^2)}{(y^3+y-y^2/2)}$, 1-y < x < 1 $f_{X\|Y}$(x|$\frac{3}{4}$) = $\frac{(x+(\frac{3}{4})^2)}{((\frac{3}{4})^3+\frac{3}{4}-(\frac{3}{4})^2/2)} = \frac{64}{57}$(x+$\frac{9}{16}$), $\frac{1}{4}$ < x < 1. P(X > $\frac{1}{2}$|Y = $\frac{3}{4}$) = $\int_{1/2}^1 f_{X\|Y}(x\|\frac{3}{4})\,dx = \int_{1/2}^1 \frac{64}{57}(x+\frac{9}{16})\,dx$ = 14/19 X and Y indep? f(x,y) = $\frac{12}{7}(x+y^2)$I(x,y), -∞ < x < ∞, -∞ < y < ∞, where I(x,y) = $\begin{cases}1, x<1, y<1, x+y>1 \\ 0, otherwise\end{cases}$ Note f(x,y) ≠ h(x)g(y), so X and Y not indep
Joint prob dist of Fn of r.v.	
Let X_1 and X_2 be indep r.v., both uniformly dist on (0,1), i.e. $f_{X_1,X_2}(x_1,x_2)$ = 1 on (0,1) Find joint pdf of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$ Find pdf of Y_1 	$y_1 = x_1 + x_2$, $y_2 = x_1 - x_2$. $x_1 = (y_1 + y_2)/2$, $x_2 = (y_1 - y_2)/2$ J(x_1, x_2) = $\begin{vmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2}\end{vmatrix} = \begin{vmatrix}1 & 1 \\ 1 & -1\end{vmatrix}$ = -2. f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)$\frac{1}{\|J(x_1,x_2)\|}$ = 1*$\frac{1}{2}$ f_{Y_1,Y_2}(y_1, y_2) = $\begin{cases}\frac{1}{2}, 0 \le y_1+y_2 \le 2, 0 \le y_1-y_2 \le 2 \\ 0, otherwise\end{cases}$ For 0 < y_1 < 1, $f_{Y_1}(y_1) = \int_{-y_1}^{y_1}\frac{1}{2}\,dy_2 = y_1$. For 1 < y_1 < 2, $f_{Y_1}(y_1) = \int_{y_1-2}^{-y_1+2}\frac{1}{2}\,dy_2 = 2 - y_1$ $f_{Y_1}(y_1) = \begin{cases} y_1, 0<y_1<1 \\ 2-y_1, 1<y_1<2 \\ 0, otherwise\end{cases}$
$f_{X_1,X_2}(x_1,x_2) = \begin{cases}e^{-(x_1+x_2)}, x_1>0, x_2>0 \\ 0, otherwise\end{cases}$ Find joint pdf of $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1+X_2}$ Find marginal pdf of Y_2	$y_1 = x_1 + x_2$ and $y_2 = \frac{x_1}{x_1+x_2}$. $x_1 = y_1y_2$, $x_2 = y_1(1-y_2)$. J(x_1, x_2) = $\begin{vmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2}\end{vmatrix} = \begin{vmatrix}1 & 1 \\ \frac{x_2}{(x_1+x_2)^2} & \frac{-x_1}{(x_1+x_2)^2}\end{vmatrix} = \frac{-1}{x_1+x_2}$ f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)$\frac{1}{\|J(x_1,x_2)\|} = e^{-y_1}\frac{1}{1/y_1} = y_1e^{-y_1}$, y_1 > 0, 0 < y_2 < 1 f_{Y_2}(y_2) = $\int_0^\infty f_{Y_1,Y_2}(y_1,y_2)\,dy_1 = \int_0^\infty y_1e^{-y_1}\,dy_1$ = 1, 0 < y_2 < 1 (by parts)
Extra	
Joint pdf of X and Y is f(x,y) = c for 0 < x < 1 and 0 < y < 2 where c is a constant and zero otherwise.	$\int_0^2\int_0^1 c\,dx\,dy$ = 1. c = 1/2. P(X > 1/2) = $\int_{1/2}^2\int_0^1 1/2\,dx\,dy$ = 3/4. P(X > Y) = $\int_0^1\int_y^1 1/2\,dx\,dy$ = 1/4
Suppose 3 balls are chosen w/o replacement from an urn consisting of 5W and 8R balls. Let X_i = 1 if i^{th} ball chosen is W and 0 otherwise. Give joint pmf of X_1, X_2 and X_1, X_2, X_3 Suppose now W balls are numbered, let Y_i = 1 if the i^{th} W ball is chosen and 0 otherwise. Find joint pmf of Y_1, Y_2 and Y_1, Y_2, Y_3.	X: p(0,0) = $\frac{8}{13}\frac{7}{12} = \frac{14}{39}$. p(0, 1) = p(1,0) = $\frac{8}{13}\frac{5}{12} = \frac{10}{39}$. p(1,1) = $\frac{5}{39}$ X: p(0,0,0) = $\frac{28}{143}$. p(0,0,1) = p(0,1,0) = p(1,0,0) = $\frac{70}{429}$. p(0,1,1) = p(1,0,1) = p(1,1,0) = $\frac{40}{429}$. p(1,1,1) = $\frac{5}{143}$ Y: p(0,0) = $\frac{\binom{2}{0}\binom{11}{3}}{\binom{13}{3}} = \frac{15}{26}$. p(1,1) = $\frac{\binom{2}{2}\binom{11}{1}}{\binom{13}{3}} = \frac{1}{26}$. p(0,1) = p(1,0) = $\frac{\binom{1}{0}\binom{1}{1}\binom{11}{2}}{\binom{13}{3}} = \frac{5}{26}$. Y: p(0,0,0) = $\frac{10}{13}\frac{9}{12}\frac{8}{11} = \frac{60}{143}$. p(0,0,1) = p(0,1,0) = p(1,0,0) = $\frac{\binom{10}{2}\binom{1}{1}}{\binom{13}{3}} = \frac{45}{286}$. p(i,j,k) = $\frac{\binom{2}{2}\binom{10}{1}}{\binom{13}{3}} = \frac{5}{143}$. p(1,1,1) = $\frac{1}{286}$
Joint pdf of X and Y is f(x,y) = $\frac{6}{7}$(x² + $\frac{xy}{2}$), 0 < x < 1, 0 < y < 2. Verify is a joint density fn. Compute density fn of X P(X>Y). P(Y > 1/2 | X < 1/2). E(X).	$\int_0^2\int_0^1\frac{6}{7}(x^2+\frac{xy}{2})\,dx\,dy$ = 1. f_X(x) = $\int_0^2\frac{6}{7}(x^2+\frac{xy}{2})\,dy = \frac{6}{7}(2x^2+x)$, 0 < x < 1 P(X > Y) = $\int_0^1\int_0^x\frac{6}{7}(x^2+\frac{xy}{2})\,dy\,dx = \frac{15}{56}$. P(Y > $\frac{1}{2}$|X < $\frac{1}{2}$) = $\frac{P(Y>1/2 \text{ and } X<1/2)}{P(X<1/2)} = \frac{\int_{1/2}^2\int_0^{1/2}\frac{6}{7}(x^2+\frac{xy}{2})dxdy}{\int_0^{1/2}\frac{6}{7}(2x^2+x)dx} = \frac{69}{80}$ E(X) = $\int xf(x)dx = \int_0^1 x\frac{6}{7}(2x^2+x)\,dx = \frac{5}{7}$.
Suppose n points are indepently chosen at random on perimeter of circle, and we want the prob that they all lie in some semicircle. Let P_1,...,P_n denote the n points. Let A = {all points are contained in the same semicircle}. A_i = {all points lie in semicircle beginning at point P_i and gg clockwise for 180°}, i = 1,...,n	Express A in terms of A_i. A = $\cup_{i=1}^n A_i$ (A is true as long as one of the A_i is true) Are A_i mutually exclusive (ME)?. Yes P(A) = P($\cup_{i=1}^n A_i$) = $\cup_{i=1}^n P(A_i) = \sum_{i=1}^n\left(\frac{1}{2}\right)^{n-1}$ (ME) = n$\left(\frac{1}{2}\right)^{n-1}$
3 pts, X_1, X_2, X_3 are selected at random on line L.	P(X_2 lies btw X_1 and X_3) = 1/3 (by symmetry). Any of the 3 pts equally likely to be middle one
2 pts are selected randomly on line of length L so as to be on opp sides of the midpoint of line. i.e. X and Y are indep r.v. and X ~ U(0,L/2), Y ~ U(L/2, L).	 f(x) = 2/L, 0 < x < L/2. f(y) = 2/L, L/2 < y < L. f(x,y) = (2/L)(2/L) = 4/L², 0 < x < L/2, L/2 < y < L P(Y - X > L/3) = $\int_0^{L/6}\int_{L/2}^L 4/L^2 dy\,dx + \int_{L/6}^L\int_{x+L/3}^L 4/L^2 dy\,dx$ = 7/9
Show f(x,y) = 1/x, 0 < y < x < 1 is a joint density fn. Assume f is joint density fn of X,Y	 $\int_0^1\int_0^x 1/x\,dy\,dx$ = 1. f_Y(y) = $\int_y^1 1/x\,dx$ = -ln(y), 0 < y < 1. f_X(x) = $\int_0^x 1/x\,dy$ = 1, 0 < x < 1. E(X) = $\int_0^1 x(1)\,dx$ = 1/2. E(Y) = $\int_0^1 y(-\ln y)\,dy$ = 1/2 (by parts)
Let f(x,y) = 24xy, 0 ≤ x ≤ 1, 0 ≤ y ≤ 1, 0 ≤ x + y ≤ 1	Show f(x,y) is a joint pdf. f(x,y) ≥ 0. AND $\int_0^1\int_0^{1-y} f(x,y)\,dx\,dy$ = 1 f_X(x) = $\int_0^{1-x} 24xy\,dy$ = 12x(1-x)², 0 < x < 1. E(X) = $\int_0^1 x(12x)(1-x)^2\,dx$ = 2/5 By symmetry, E(Y) = E(X) = 2/5
Number of ppl that enter a store in a given hour is a Poisson r.v. w λ = 10. Compute conditional prob that at most 3 men enter store, given that 10 women entered in the hour. Assumptions?	Let X = num of men who entered, Y = num of women who entered. X + Y ~ Poisson(10) If we assume prob of men entering p = 1/2, women entering is 1-p, and X and Y are indep, then X ~ Poisson(1x0*1/2 = 5). Y ~ Poisson(5). P(X ≤ 3|Y = 10) = P(X ≤ 3) (indep) = $e^{-5} + \frac{e^{-5}5^1}{1!} + \frac{e^{-5}5^2}{2!} + \frac{e^{-5}5^3}{3!}$
Joint density of X and Y, f(x,y) = $\begin{cases}xe^{-(x+y)}, x>0, y>0 \\ 0, otherwise\end{cases}$	f_X(x) = $\int_0^\infty xe^{-(x+y)}dy$ = $xe^{-x}\int_0^\infty e^{-y}dy$ = xe^{-x}, x > 0 f_Y(y) = $\int_0^\infty xe^{-(x+y)}dx$ = $e^{-y}\int_0^\infty xe^{-x}dx$ = e^{-y}, y > 0. ∵ f(x,y) = $xe^{-(x+y)}$ = $f_X(x)f_Y(y)$. X and Y indep

Are X and Y indep? What if f(x,y) = $\begin{cases} 2, 0<x<y, 0<y<1 \\ 0, otherwise \end{cases}$	$f_X(x) = \int_x^1 2\,dy = 2(1-x)$, 0 < x < 1. $f_Y(y) = \int_0^y 2\,dx = 2y$, 0 < y < 1 $f_X(x)f_Y(y) = 4y(1-x) \neq 2$ = f(x,y). X and Y not indep
Joint density fn of X and Y is f(x,y) = $\begin{cases} x+y, 0<x<1, 0<y<1 \\ 0, otherwise \end{cases}$	Are X and Y indep? No, cause f(x,y) cannot be written in the form f(x,y) = h(x)g(y) Find density fn of X. $f_X(x) = \int_0^1 x+y\,dy$ = x + 1/2, 0 < x < 1 P(X+Y < 1) = $\int_0^1\int_0^{1-x} x+y\,dy\,dx$ = 1/3
Consider indep trials each of which result in outcome i, i = 0,1,...,k w prob $p_i = \sum_{i=0}^k p_i = 1$. Let N denote num of trials needed to obtain outcome that is not equal to 0, and let X be that outcome.	P(N = n), n ≥ 1 = $p_0^{n-1}(1-p_0)$. P(X = j), j = 1,...,k = $\frac{p_j}{1-p_0}$ Show P(N=n, X=j) = (outcome 0)$^{n-1}$*(outcome j) = $p_0^{n-1}p_j = p_0^{n-1}(1-p_0)\frac{p_j}{1-p_0}$ = P(N=n)P(X=j)
Weekly sales at a restaurant is a normal r.v. w mean \$2200 and s.d. \$230. Let W = $X_1 + X_2$, where $X_1 \sim N(2200, 230^2)$, $X_2 \sim N(2200, 230^2)$	P(2 weeks sales > 5000)? W ~ N(4400, 2(230)^2). P(W > 5000) = $P(\frac{W-4400}{\sqrt{2(230)^2}} > \frac{5000-4400}{\sqrt{2(230)^2}})$ = P(Z > 1.8446) = 0.0326 P(weekly sales > 2000 in at least 2 of the next 3 weeks)? p = P(X > 2000) = $P(\frac{X-2200}{230} > \frac{2000-2200}{230})$ = P(Z > -0.87) = 0.8078. 3 weeks + 2weeks = $p^3 + 3p^2(1-p)$
2 dice are rolled. Let X = largest value and Y = smallest value. Compute conditional mass fn of Y given X = i, for i = 1,2,...,6. Are X and Y indep? X and Y are not indep. In particular, Y ≤ X	For j = i, P(Y = i|X = i) = $\frac{P(Y=i,X=i)}{P(X=i)} = \frac{1/36}{P(X=i)}$. For j < i, P(Y = j|X = i) = $\frac{P(Y=j,X=i)}{P(X=i)} = \frac{2/36}{P(X=i)}$ For a fixed i, 1 = $\sum_{j=1}^{i-1} P(Y=j\|X=i)$ + P(Y = i|X = i) = $\sum_{j=1}^{i-1}\frac{2/36}{P(X=i)} + \frac{1/36}{P(X=i)}$. P(X = i) = $\sum_{j=1}^{i-1} 2/36 + 1/36$. (Multiply P(X = i) on both sides). P(X = i) = (i-1)(2/36) + 1/36 = (2i-1)/36 P(Y = j|X = i) = $\begin{cases} 1/(2i-1), j=i \\ 2/(2i-1), j<i \end{cases}$
Joint density fn of X and Y, f(x,y) = $xe^{-x(y+1)}$, x > 0, y > 0 Find conditional density of X, given Y = y and that of Y, given X = x Find density fn of Z = XY	$f_Y(y) = \int_0^\infty xe^{-x(y+1)}dx = \frac{1}{(y+1)^2}$, y > 0. $f_{X\|Y}(x\|y) = \frac{f(x,y)}{f_Y(y)} = \frac{xe^{-x(y+1)}}{1/(y+1)^2} = (y+1)^2xe^{-x(y+1)}$, x > 0 $f_X(x) = \int_0^\infty xe^{-x(y+1)}dy = e^{-x}$, x > 0. $f_{Y\|X}(y\|x) = \frac{f(x,y)}{f_X(x)} = \frac{xe^{-x(y+1)}}{e^{-x}} = xe^{-xy}$, y > 0 $F_Z(z) = P(Z \le z) = P(XY \le z) = \int_0^\infty\int_0^{z/x} xe^{-x(y+1)}\,dy\,dx = 1 - e^{-z}$, z > 0. $f_Z(z) = \frac{d}{dz}F_Z(z) = e^{-z}$, z > 0
Joint density fn of X and Y, f(x,y) = $c(x^2 - y^2)e^{-x}$, 0 ≤ x < ∞, -x ≤ y ≤ x Find conditional dist of Y given X = x 	$f_X(x) = \int_{-x}^x c(x^2-y^2)e^{-x}dy = ce^{-x}x^3(\frac{4}{3})$. $f_{Y\|X}(y\|x) = \frac{f(x,y)}{f_X(x)} = \frac{c(x^2-y^2)e^{-x}}{ce^{-x}x^3(4/3)} = \frac{3}{4}\frac{(x^2-y^2)}{x^3}$, -x ≤ y ≤ x $F_{Y\|X}(a\|x) = \int_{-\infty}^a f_{Y\|X}(y\|x)dy = \int_{-x}^a \frac{3}{4}\frac{(x^2-y^2)}{x^3}dy = \frac{3}{4x^3}(x^2a - \frac{a^3}{3} + \frac{2a^3}{3})$ $F_{Y\|X}(y\|x) = \frac{3}{4x^3}(x^2y - \frac{y^3}{3} + \frac{2y^3}{3})$, -x < y < x
If X and Y have joint density fn f(x,y) = $\frac{1}{x^2y^2}$, x ≥ 1, y ≥ 1 Compute joint density fn of U = XY, V = X/Y What are the marginal densities? 	u = xy, v = x/y. y = $\sqrt{u/v}$. x = $\sqrt{uv}$. J(x, y) = $\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} y & x \\ \frac{1}{y} & -\frac{x}{y^2} \end{vmatrix} = y(-\frac{x}{y^2}) - x(\frac{1}{y}) = -\frac{2x}{y}$. |J(x,y)| = $\frac{2x}{y}$ = 2v $f_{U,V}(u,v) = f_{X,Y}(x,y)\frac{1}{\|J(x,y)\|} = \frac{1}{x^2y^2}\frac{y}{2x} = \frac{1}{(uv)(u/v)}\frac{1}{2v} = \frac{1}{2vu^2}$, u ≥ v, uv ≥ 1 (since $\sqrt{u/v} \ge 1$ and $\sqrt{uv} \ge 1$) $f_U(u) = \int_{1/u}^u \frac{1}{2vu^2}dv = \frac{1}{u^2}\ln u$, u ≥ 1 For v > 1, $f_V(v) = \int_v^\infty \frac{1}{2vu^2}du = \frac{1}{2v^2}$, v > 1 For v < 1, $f_V(v) = \int_{1/v}^\infty \frac{1}{2vu^2}du = \frac{1}{2}$, 0 < v < 1
If X_1 and X_2 are indep exponential r.v. each w parameter λ, find the joint density fn of $Y_1 = X_1 + X_2$ and $Y_2 = e^{X_1}$	$y_1 = x_1 + x_2$. $y_2 = e^{x_1}$. So $x_1 = \ln y_2$. $x_2 = y_1 - \ln y_2$. J(x_1, x_2) = $\begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ e^{x_1} & 0 \end{vmatrix} = -e^{x_1} = -y_2$ f(x_1, x_2) = $\lambda e^{-\lambda x_1}\lambda e^{-\lambda x_2}$ (since indep), x_1 > 0, x_2 > 0 $f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(x_1,x_2)\frac{1}{\|J(x_1,x_2)\|} = \lambda^2 e^{-\lambda y_1}\frac{1}{y_2}$, y_2 > 1, $y_1 > \ln y_2$
Suppose X and Y are indep geometric r.v. w param p. Without any computations, what do you think is value of P(X = i|X + Y = n)?	Given 2nd success occur at nth trial, 1st success can occur at any of 1st n-1 trials w prob 1/(n-1) P(X = i|X + Y = n) = $\frac{P(X=i,X+Y=n)}{P(X+Y=n)} = \frac{P(X=i,Y=n-i)}{P(X+Y=n)} = \frac{P(X=i)P(Y=n-i)}{P(X+Y=n)} = \frac{p(1-p)^{i-1}*p(1-p)^{n-i-1}}{\binom{n-1}{2-1}p^2(1-p)^{n-2}} = \frac{1}{n-1}$
If X is exponential w rate λ, find P{[X] = n, X - [X] ≤ x}, where [x] is defined as largest integer ≤ x. Can you conclude that [X] and X are indep?	P{[X] = n, X - [X] ≤ x} = P(n < X < n+x) = $\int_n^{n+x}\lambda e^{-\lambda x}\,dx = e^{-\lambda n}(1-e^{-\lambda x})$ So indep

Expectation of Sums of r.v.	
P(a ≤ X ≤ b) ⇒ a ≤ E(X) ≤ b	Proof (cts case). $\int_{-\infty}^\infty af(x)\,dx \le \int_{-\infty}^\infty xf(x)\,dx \le \int_{-\infty}^\infty bf(x)\,dx$. $a\int_{-\infty}^\infty f(x)\,dx$ ≤ E(X) ≤ $b\int_{-\infty}^\infty f(x)\,dx$. a ≤ E(X) ≤ b
	f(x,y) = $\begin{cases} 2, x>0, y>0, x+y<1 \\ 0, otherwise \end{cases}$, find E(XY) E(XY) = $\int_0^1\int_0^{1-x} xy(2)\,dy\,dx = \int_0^1 x\int_0^{1-x} 2y\,dy\,dx = \int_0^1 x[y^2\|_0^{1-x}\,dx$ = ... = 1/12
If E(X) & E(Y) are finite, E(X + Y) = E(X) + E(Y)	Proof (cts case). E(X+Y) = $\int_{-\infty}^\infty\int_{-\infty}^\infty (x+y)f(x,y)\,dx\,dy = \int_{-\infty}^\infty\int_{-\infty}^\infty xf(x,y)\,dy\,dx + \int_{-\infty}^\infty\int_{-\infty}^\infty yf(x,y)\,dx\,dy$ = $\int_{-\infty}^\infty x\int_{-\infty}^\infty f(x,y)\,dy\,dx + \int_{-\infty}^\infty y\int_{-\infty}^\infty f(x,y)\,dx\,dy = \int_{-\infty}^\infty xf_X(x)\,dx + \int_{-\infty}^\infty yf_Y(y)\,dy$ = E(X) + E(Y)
Let X_1,..., X_n be indep and identically distributed r.v. having dist F(x) and E(X_i) = μ If $\bar{X}$ (sample mean) = $\frac{1}{n}\sum_{i=1}^n X_i \Rightarrow E(\bar{X}) = \mu$	E($\bar{X}$) = $E(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n}E(\sum_{i=1}^n X_i) = \frac{1}{n}\sum_{i=1}^n E(X_i) = \frac{1}{n}\sum_{i=1}^n \mu = \frac{1}{n}(n\mu) = \mu$
If X~Binomial(n,p), then E(X) = np	X = $X_1 + X_2 + ... + X_n$ where X_i = $\begin{cases} 1, if\ ith\ trial\ is\ success \\ 0, if\ ith\ trial\ is\ failure \end{cases}$. E($X_i$) = 1P($X_i$ = 1) + 0P(X_i = 0) = 1p = p E(X) = E(X_1) + ... + E(X_n) = np
Negative Binomial: If X is num of trials until total r successes obtained, E(X) = r/p	X = $X_1 + X_2 + ... + X_r$, X_i = num of trials until next success ~ Geometric(p). E(X_i) = 1/p E(X) = E(X_1) + E(X_2) + ... + E(X_r) = r/p
Hypergeometric: Take n balls from urn containing m W and N-m B balls, E(Num of white balls selected) = nm/N	Let X = num of W balls selected, Y_i = $\begin{cases} 1, if\ ith\ ball\ selected\ is\ W \\ 0, otherwise \end{cases}$. X = $Y_1 + Y_2 + ... + Y_n$. E(Y_i) = 1P(Y_i = 1) = m/N. E(X) = E(Y_1) + ... + E(Y_n) = nm/N

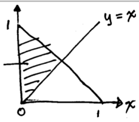
Hat throwing: E(num of ppl that select their own hat) = 1	Let X = num of matches, $X_i = \begin{cases} 1, if\ ith\ person\ select\ own\ hat \\ 0, otherwise \end{cases}$. $X = X_1 + \dots + X_N$. $P(X_i = 1) = 1/N$ $E(X_i) = 1P(X_i = 1) = 1/N$. $E(X) = N(1/N) = 1$
Coupon problem: E(num of coupons to be collected to obtain complete set) = $1 + \frac{N}{N-1} + \frac{N}{N-2} + \dots + \frac{N}{1}$	Let X = num of coupons for complete set, X_i = num of additional coupons after i distinct types collected in order to obtain another distinct type, i = 0,1,2,...,N-1. $X = X_0 + X_1 + \dots + X_{N-1}$ $X_0 = 1$. $X_1 \sim$ Geometric($\frac{N-1}{N}$), $X_2 \sim$ Geometric($\frac{N-2}{N}$),... , $X_{N-1} \sim$ Geometric($\frac{1}{N}$). $E(X) = E(X_0) + \dots + E(X_{N-1}) = 1 + \frac{1}{(N-1)/N} + \frac{1}{(N-2)/N} + \dots + \frac{1}{1/N} = 1 + \frac{N}{N-1} + \frac{N}{N-2} + \dots + \frac{N}{1} = N(\frac{1}{N} + \frac{1}{N-1} + \frac{1}{N-2} + \dots + 1)$

Covariance, Variance of Sums, Correlations

X and Y indep ⇒ E[g(X)h(Y)] = E[g(X)]E[h(Y)]	Proof (cts case). $E[g(X)h(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)\,dx\,dy$ (since indep) = $\int_{-\infty}^{\infty} g(x)f_X(x)dx * \int_{-\infty}^{\infty} h(y)f_Y(y)\,dy$ = E[g(X)]E[h(Y)]
Cov(X, Y) = E(XY) - E(X)E(Y)	Proof. Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY - XE(Y) - YE(X) + E(X)E(Y)] = E(XY) - E[XE(Y)] - E[YE(X)] + E[E(X)E(Y)] = E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y) = E(XY) - E(X)E(Y)
$Cov(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j) = \sum_{i=1}^n\sum_{j=1}^m Cov(X_i, Y_j)$	Proof. Let $E(X_i) = \mu_i$, i = 1,...,n and $E(Y_j) = \nu_j$, j = 1,...,m. $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \mu_i$. $E[\sum_{j=1}^m Y_j] = \sum_{j=1}^m E(Y_j) = \sum_{j=1}^m \nu_j$. $Cov(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j) = E[(\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i)(\sum_{j=1}^m Y_j - \sum_{j=1}^m \nu_j)] = E[\sum_{i=1}^n (X_i - \mu_i)\sum_{j=1}^m (Y_j - \nu_j)] =$ $E[\sum_{i=1}^n\sum_{j=1}^m (X_i - \mu_i)(Y_j - \nu_j)] = \sum_{i=1}^n\sum_{j=1}^m E[(X_i - \mu_i)(Y_j - \nu_j)] = \sum_{i=1}^n\sum_{j=1}^m Cov(X_i, Y_j)$
$Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, Y_j)$	Proof. $Var(\sum_{i=1}^n X_i) = Cov(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j)$ (ii) $= \sum_{i=1}^n\sum_{j=1}^n Cov(X_i, X_j)$ (iv) $= \sum_{i=1}^n Cov(X_i, X_i) + \sum\sum_{i\neq j} Cov(X_i, X_j) = \sum_{i=1}^n Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, Y_j)$ (ii + i)
If X~Binomial(n,p) then Var(X) = np(1-p)	Let X = num of success in n indep trials. $X_i = \begin{cases} 1, if\ ith\ trial\ is\ success \\ 0, if\ ith\ trial\ is\ failure \end{cases}$. $X = X_1 + \dots + X_n$ and X_i's are indep $E(X_i) = 1P(X_i = 1) = p$. $E(X_i^2) = 1^2P(X_i = 1) = p$. $Var(X_i) = E(X_i^2) - [E(X)]^2 = p - p^2 = p(1-p)$ $Var(X) = Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n)$ (indep) = np(1-p)
Let $X_1,\dots,X_n$ be indep and identically distributed r.v. each having expected value μ and var σ^2. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ s^2 is an unbiased estimator of σ^2	$Var(\overline{X}) = Var(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n^2}Var(\sum_{i=1}^n X_i) = \frac{1}{n^2}\sum_{i=1}^n Var(X_i)$ (since X_i indep) $= \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$ $(n-1)s^2 = \sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n (X_i - \mu + \mu - \overline{X})^2 = \sum_{i=1}^n [(X_i - \mu)^2 + (\mu - \overline{X})^2 - 2(X_i - \mu)(\mu - \overline{X})] =$ $\sum_{i=1}^n (X_i - \mu)^2 + \sum_{i=1}^n (\mu - \overline{X})^2 - 2(\overline{X} - \mu)\sum_{i=1}^n (X_i - \mu) = \sum_{i=1}^n (X_i - \mu)^2 + n(\mu - \overline{X})^2 - 2(\overline{X} - \mu)(\sum_{i=1}^n X_i - n\mu) =$ $\sum_{i=1}^n (X_i - \mu)^2 + n(\mu - \overline{X})^2 - 2(\overline{X} - \mu)(n\overline{X} - n\mu) = \sum_{i=1}^n (X_i - \mu)^2 + n(\mu - \overline{X})^2 - 2n(\overline{X} - \mu)(\overline{X} - \mu) =$ $\sum_{i=1}^n (X_i - \mu)^2 - n(\mu - \overline{X})^2$ $E[(n-1)s^2] = E[\sum_{i=1}^n (X_i - \mu)^2 - n(\mu - \overline{X})^2]$. $(n-1)E(s^2) = E[\sum_{i=1}^n (X_i - \mu)^2] - E[n(\mu - \overline{X})^2] = \sum_{i=1}^n E[(X_i - \mu)^2] -$ $nE[(\mu - \overline{X})^2] = \sum_{i=1}^n \sigma^2 - nVar(\overline{X}) = n\sigma^2 - n(\frac{\sigma^2}{n}) = (n-1)\sigma^2$. So $E(s^2) = \sigma^2$.
	$f(x,y) = \begin{cases} 2, x > 0, y > 0, x + y < 1 \\ 0, otherwise \end{cases}$, find Cov(X,Y) $E(X) = \int_0^1\int_0^{1-x} x(2)\,dy\,dx = 1/3$. $E(Y) = \int_0^1\int_0^{1-x} y(2)\,dy\,dx = 1/3$. $E(XY) = \int_0^1\int_0^{1-x} xy(2)\,dy\,dx = 1/12$ So Cov(X, Y) = E(XY) - E(X)E(Y) = 1/12 - (1/3)(1/3) = -1/36 < 0. As x incr, y decr
$-1 \le \rho(X, Y) \le 1$ Note $\rho(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$	Proof. Suppose $Var(X) = \sigma_X^2$, $Var(Y) = \sigma_Y^2$. $0 \le Var(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}) = Var(\frac{X}{\sigma_X}) + Var(\frac{Y}{\sigma_Y}) + 2Cov(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}) = \frac{1}{\sigma_X^2}Var(X) + \frac{1}{\sigma_Y^2}Var(Y) + \frac{2}{\sigma_X\sigma_Y}Cov(X, Y) = 1 + 1 + 2\rho(X, Y) = 2[1 + \rho(X, Y)]$. $\rho(X, Y) \ge -1$. Similarly, starting from $0 \le Var(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y})$, $\rho(X, Y) \le 1$
Let I_A and I_b be indicator variables, $I_A = \begin{cases} 1, if\ A\ occurs \\ 0, otherwise \end{cases}$, $I_B = \begin{cases} 1, if\ B\ occurs \\ 0, otherwise \end{cases}$	$E(I_A) = 1P(A) + 0P(A^C) = P(A)$. $E(I_B) = P(B)$. $I_AI_B = \begin{cases} 1, if\ A\ and\ B\ occurs \\ 0, otherwise \end{cases}$. $E(I_AI_B) = P(AB)$ $Cov(I_A, I_B) = E(I_AI_B) - E(I_A)E(I_B) = P(AB) - P(A)P(B) = P(B)[\frac{P(AB)}{P(B)} - P(A)] = P(B)[P(A\mid B) - P(B)]$, i.e.cov = 0 if A indep of B
Let $X_1, \dots, X_n$ be indep and identically distributed r.v. w variance σ^2. Show $Cov(X_i - \overline{X}, \overline{X}) = 0$	$Cov(X_i + (-\overline{X}), \overline{X}) = Cov(X_i, \overline{X}) + Cov(-\overline{X}, \overline{X})$ (iv) $= Cov(X_i, \overline{X}) - Cov(\overline{X}, \overline{X}) = Cov(X_i, \frac{1}{n}\sum_{j=1}^n X_j) - Var(\overline{X}) =$ $\frac{1}{n}\sum_{j=1}^n Cov(X_i, X_j) - Var(\overline{X})$ (iv) $= \frac{1}{n}Cov(X_i, X_i) - \frac{\sigma^2}{n}$ (since X_i, X_j indep, cov = 0) $= \frac{1}{n}Var(X_i) - \frac{\sigma^2}{n} = \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0$

Conditional Expectation

X,Y ~ Geometric(p) and indep, find $E(X\mid X+Y = n)$	$P(X = i\mid X+Y = n) = \frac{1}{n-1}$, i = 1,2,...,n-1. $E(X\mid X+Y = n) = \sum_{i=1}^{n-1} iP(X = i\mid X + Y = n) = \sum_{i=1}^{n-1} i\frac{1}{n-1} = \frac{1}{n-1}\frac{(n-1)n}{2} = \frac{n}{2}$
$f(x,y) = \begin{cases} 2, x > 0, y > 0, x + y < 1 \\ 0, otherwise \end{cases}$, find $E(X\mid Y = y)$	$f_Y(y) = \int_0^{1-y} f(x,y)\,dx = \int_0^{1-y} 2\,dx = 2(1-y)$, 0 <y< 1. $f_{X\mid Y}(x\mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y}$, 0 < x < 1-y. So $X\mid Y = y \sim U(0, 1-y)$ $E(X\mid Y = y) = \int_0^{1-y} x\frac{1}{1-y}dx = \frac{1-y}{2}$, 0 < y < 1
Toss 2 dice. X, Y = largest, smallest value. find $E(Y\mid X = 4)$	P(Y = 1, X = 4) = 2/7. $P(Y = 2\mid X=4) = 2/7 = P(Y=3\mid X=4)$. $P(Y=4\mid X=4) = 1/7$ $E(Y\mid X=4)$ = 1(2/7) + 2(2/7) + 3(2/7) + 4(1/7) = 16/7
$E(X) = \sum_y E(X\mid Y = y)P(Y = y)$	Proof (discrete case). $\sum_y E(X\mid Y = y)P(Y = y) = \sum_y\sum_x xP(X = x\mid Y = y)\,P(Y = y) =$ $\sum_y\sum_x x\frac{P(X=x,Y=y)}{P(y=y)}P(Y = y) = \sum_y\sum_x x\,P(X = x, Y = y) = \sum_x x\sum_y P(X = x, Y = y) = \sum_x xP(X = x)$ = E(X)
Miner and 3 doors. 1st door lead to exit after 3 hrs. 2nd door lead to starting place after 5 hrs. 3rd door lead to starting place after 7 hours. Assume miner at all times equally likely to choose any door, expected time until exit?	Let X = amt of time until exit, Y = door chosen $E(X) = \sum_{y=1}^3 E(X\mid Y = y)P(Y = y) = E(X\mid Y=1)P(Y=1) + E(X\mid Y=2)P(Y=2) + E(X\mid Y=3)P(Y=3) =$ 3(1/3) + [5+E(X)](1/3) + [7+E(X)](1/3). So E(X) = 15
$f(x,y) = \begin{cases} 2, x > 0, y > 0, x + y < 1 \\ 0, otherwise \end{cases}$, 3 diff ways of finding E(X)	1. $E(X) = \int_0^1\int_0^{1-y} xf(x,y)\,dx\,dy = 1/3$. $\{E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy\}$ 2. $f_X(x) = \int_0^{1-x} 2\,dy = 2(1-x)$, 0 < x < 1. $E(X) = \int_0^1 x(2)(1-x)\,dx = 1/3$. $\{E(X) = \int_{-\infty}^{\infty} xf(x)dx\}$ 3. $f_Y(y) = 2(1-y)$, 0 < y < 1. $E(X\mid Y = y) = \frac{1-y}{2}$. $E(X) = \int_0^1 \frac{1-y}{2}2(1-y)\,dy = 1/3$. $\{E(X) = \int_{-\infty}^{\infty} E(X\mid Y = y)f_Y(y)\,dy\}$
Suppose num of ppl entering store on a given day is r.v. w mean 50. Suppose amt of money spent by customers is are indep r.v. w mean \$8. Assume amt of money spent by customer is indep of total num of customers in store. Find expected amt of money spent in store on a given day.	Let N = num of customers that enter store, X_i = amt spend by i[th] customer Total amt of money spent = $\sum_{i=1}^N X_i$ which is a r.v. $E(\sum_{i=1}^N X_i) = E(E(\sum_{i=1}^N X_i\mid N)) = \sum_{n=0}^{\infty}[E(\sum_{i=1}^N X_i\mid N = n)P(N = n)]$ (cond expectation over all n) $= \sum_{n=0}^{\infty}[E(\sum_{i=1}^n X_i\mid N = n)P(N = n)]$ (change of var) $= \sum_{n=0}^{\infty}[E(\sum_{i=1}^n X_i)P(N = n)]$ (N and X_i are indep) $= \sum_{n=0}^{\infty}[(\sum_{i=1}^n E(X_i))\,P(N = n)] = \sum_{n=0}^{\infty}[nE(X_1)P(N = n)]$ ($E(X_i)$ all same) = $E(X_1)\sum_{n=0}^{\infty}[nP(N = n)] = E(X_1)E(N) = 8 * 50 = 400$
$P(E) = \sum_y P(E\mid Y = y)P(Y = y)$ $P(E) = \int_{-\infty}^{\infty} P(E\mid Y = y)f_Y(y)\,dy$	Proof. Let E denote any arbitrary event and $X = \begin{cases} 1, if\ E\ occurs \\ 0, otherwise \end{cases}$. $E(X) = 1P(E) + 0P(E^C) = P(E)$ $E(X\mid Y=y) = 1P(X=1\mid Y=y) + 0P(X=0\mid Y=y) = P(X=1\mid Y=y) = P(E\mid Y=y)$

	E(X) = $\begin{cases} \sum_y E(X\|Y=y)P(Y=y), if\ Y\ discrete \\ \int_{-\infty}^{\infty} E(X\|Y=y)f_Y(y)\,dy, if\ Y\ cts \end{cases}$. P(E) = E(X) = $\begin{cases} \sum_y P(E\|Y=y)P(Y=y), if\ Y\ discrete \\ \int_{-\infty}^{\infty} P(E\|Y=y)f_Y(y)\,dy, if\ Y\ cts \end{cases}$
f(x,y) = $\begin{cases} 2, x>0, y>0, x+y<1 \\ 0, otherwise \end{cases}$, find P(X < Y)  OR $\int_0^{1/2}\int_x^{1-x} 2\,dy\,dx$ = 1/2	P(X < Y) = $\int_0^1 P(X<Y\|Y=y)f_Y(y)\,dy$, where $f_Y(y)$ = 2(1-y), 0 < y < 1. X|Y = y ~ U(0, 1-y) (pg 5), $f_{X\|Y}(x\|y)$ = $\frac{1}{1-y}$, 0 < x < 1-y For y < 1/2: P(X < y|Y = y) = $\int_0^y \frac{1}{1-y}dx = \frac{y}{1-y}$. For y > 1/2. P(X < y|Y = y) = 1 (since x + y < 1) P(X < Y) = $\int_0^{1/2}\frac{y}{1-y}2(1-y)\,dy + \int_{1/2}^1 1(2)(1-y)\,dy$ = 1/2
Suppose by any time t, num of ppl that have arrived at a train station is a Poisson r.v. w mean λt. If train arrives at station at time (indep of ppl arrival) uniformly dist over (0,T). What is mean and var of num of ppl entering train?	Let N(t) = num of arrivals by time t, Y = time train arrives E(N(Y)|Y=t) = E(N(t)|Y=t) = E(N(t)) (since N(t) and Y are indep) = λt. E(N(Y)|Y) = λY, a r.v.. E(E(N(Y)|Y)) = E(λY) = λE(Y) = λ(T/2) Var(N(Y)) = E[Var(N(Y)|Y)] + Var[E(N(Y)|Y)] Var(N(Y)|Y=t) = Var(N(t)|Y=t) = Var(N(t)) (N(t) and Y are indep) = λt Var(N(Y)|Y) = λY, a r.v.. E[Var(N(Y)|Y)] = E[λY] = λE(Y) = λ(T/2) Var[E(N(Y)|Y)] = Var(λY) = λ^2Var(Y) = λ^2(T^2/12). Thus, Var(N(Y)) = λ(T/2) + λ^2(T^2/12)
Moment Generating Functions	
$M^n(t)$ = E($X^n e^{tX}$), n ≥ 1	Proof. M(t) = E(e^{tX}). M'(t) = $\frac{d}{dt}$E(e^{tX}) = E($\frac{d}{dt}e^{tX}$) = E(Xe^{tX}). M'(0) = E(X) M''(t) = $\frac{d}{dt}$E(Xe^{tX}) = E($\frac{d}{dt}Xe^{tX}$) = E(X^2e^{tX}). M''(0) = E(X^2)
X~Binomial(n,p), M(t) = $(pe^t + (1-p))^n$	M(t) = E(e^{tX}) = $\sum_{k=0}^n e^{tX}P(X=k) = \sum_{k=0}^n e^{tk}\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k}(pe^t)^k(1-p)^{n-k}$ = $(pe^t + (1-p))^n$ (binomial expansion). M'(t) = $n(pe^t + 1-p)^{n-1}$. M'(0) = E(X) = np. M''(t) = $n(n-1)(pe^t + 1-p)^{n-2}(pe^t)^2 + n(pe^t + 1-p)^{n-1}pe^t$. M''(0) = E($X^2$) = n(n-1)$p^2$ + np. Var(X) = E(X^2) - $[E(X)]^2$ = n(n-1)p^2 + np - $[np]^2$ = np(1-p)
X~Poisson(λ), M(t) = exp[λ(e^t - 1)]	M(t) = E(e^{tX}) = $\sum_{n=0}^\infty e^{tn}P(X=n) = \sum_{n=0}^\infty e^{tn}\frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^\infty\frac{(\lambda e^t)^n}{n!} = e^{-\lambda}\sum_{n=0}^\infty\frac{(\lambda e^t)^n}{n!} = e^{-\lambda}e^{\lambda e^t}$ (expansion of e^x) = $e^{\lambda(e^t-1)}$ = exp[$\lambda(e^t-1)$]. M'(0) = E(X) = λ. M''(0) = E(X^2) = λ^2 + λ. Var(X) = λ
X~Exp(λ), M(t) = λ/(λ - t)	M(t) = E(e^{tX}) = $\int_0^\infty e^{tx}f(x)\,dx = \int_0^\infty e^{tx}\lambda e^{-\lambda x}\,dx = \lambda\int_0^\infty e^{(t-\lambda)x}\,dx = \frac{\lambda}{t-\lambda}\left[e^{(t-\lambda)x}\right\|_0^\infty$ for t < λ = $\frac{\lambda}{\lambda-t}$ M'(0) = E(X) = $\frac{1}{\lambda}$. M''(0) = E(X^2) = $\frac{2}{\lambda^2}$. Var(X) = $\frac{1}{\lambda^2}$
X~Normal(0,1), M(t) = $e^{t^2/2}$	M(t) = E(e^{tX}) = $\int_{-\infty}^\infty e^{tx}f(x)\,dx = \int_{-\infty}^\infty e^{tx}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{1}{2}(x^2-2tx)}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{1}{2}[(x-t)^2-t^2)]}\,dx$ (complete the sq) = $\frac{e^{t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{1}{2}(x-t)^2}\,dx = e^{t^2/2}\int_{-\infty}^\infty\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2}\,dx = e^{t^2/2}(1)$ (since $\int_{-\infty}^\infty\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2}$ = pdf of N(t,1)) M'(0) = E(X) = 0. M''(0) = E(X^2) = 1. Var(X) = 1
X and Y indep ⇒ $M_{X+Y}(t) = M_X(t)M_Y(t)$	Proof. $M_{X+Y}(t)$ = E[$e^{t(X+Y)}$] = E[$e^{tX}e^{tY}$] = E(e^{tX})E(e^{tY}) (since X and Y indep) = $M_X(t)M_Y(t)$
If X and Y are indep r.v., X~Binomial(n,p), Y~Binomial(m,p), what is distribution of X + Y?	$M_X(t) = (pe^t + (1-p))^n$. $M_Y(t) = (pe^t + (1-p))^m$ Since X and Y indep, $M_{X+Y}(t) = M_X(t)M_Y(t) = (pe^t + (1-p))^n(pe^t + (1-p))^m = (pe^t + (1-p))^{n+m}$ Looking at the mgf, X + Y have dist Binomial(n+m, p)
If X = (μ, σ^2) find mgf of X	Z = $\frac{X-\mu}{\sigma}$ ~ N(0,1). $M_Z(t) = e^{t^2/2}$ = E(e^{tZ}) = E($e^{t\frac{X-\mu}{\sigma}}$) = E($e^{\frac{t}{\sigma}X}e^{-\frac{t}{\sigma}\mu}$) = $e^{-\frac{t}{\sigma}\mu}$ E($e^{\frac{t}{\sigma}X}$) = (let s = $\frac{t}{\sigma}$) = $e^{-s\mu}$ E(e^{sX}). $e^{-s\mu}$ E(e^{sX}) = $e^{s^2\sigma^2/2}$. E(e^{sX}) = $e^{s^2\sigma^2/2\,+\,s\mu}$. $M_X(t) = e^{t^2\sigma^2/2\,+\,t\mu}$
Extra	
Joint pdf of X and Y, f(x,y) = 2/3 for 0 < x < 1, 0 < y < 2, x < y and 0 otherwise.	E(X) = $\int_0^1\int_x^2 x(2/3)\,dy\,dx$ = 4/9. [E(g(x,y) = $\int_{-\infty}^\infty\int_{-\infty}^\infty g(x,y)f(x,y)\,dx\,dy$] E(XY) = $\int_0^1\int_x^2 xy(2/3)\,dy\,dx$ = 7/12 E(Y) = $\int_0^1\int_x^2 y(2/3)\,dy\,dx$ = 11/9. Cov(X, Y) = E(XY) - E(X)E(Y) = 13/324 $f_X(x) = \int_x^2 2/3\,dy$ = (4-2x)/3. $f_{Y\|X}(y\|x)$ = f(x,y)/$f_X(x)$ = (2/3)/[(4-2x)/3] = 1/(2-x), x < y < 2 Then Y|X = x ~ U(x, 2). E(Y|X = x) = (x+2)/2. OR E(Y|X = x) = $\int_x^2 y(1/2-x)\,dy$ = (x+2)/2 Y|X = x ~ U(x, 2). Var(Y|X = x) = $(2-x)^2$/12
Z ~ N(0,1). $M_Z(t) = e^{t^2/2}$. Find dist of -Z	$M_{-Z}(t)$ = E($e^{t(-Z)}$) = E($e^{(-t)Z}$) = $M_Z(-t)$ = $e^{(-t)^2/2} = e^{t^2/2}$. Thus, -Z ~ N(0,1)
Discrete r.v. X has pmf P(X = -1) = 1/4, P(X = 0) = 1/2, P(X = 1) = 1/4. Fing mgf of X.	M(t) = $\sum_x e^{tx}p(x)$ = e^{-t}P(X = -1) + e^{0t}P(X = 0) = e^tP(X = 1) = (e^{2t} + 2e^t + 1)/(4e^t)
Hospital is located at center of a square w length 3 miles. If accident occur within square, then hospital sends out an ambulance. The road network is rectangular, so the travel dist from hospital, whose coordinates are (0,0) to the point (x,y) is |x| + |y|. If an accident occurs at a pt that is uniformly distributed in the sq, find the expected travel dist of the ambulance.	joint density (X, Y) at which accident occurs is f(x,y) = 1/9, -3/2 < x, y < 3/2 = f(x)f(y) where f(a) = 1/3, -3/2 < a < 3/2. Hence X and Y are indep and uniformly distributed on (-3/2, 3/2). E(|X| + |Y|) = $2\int_{-3/2}^{3/2} x(1/3)\,dx$ = 3/2
Suppose A and B each randomly and independently choose 3 out of 10 objects. Find the expected num of objects a) chosen by both A and B b) not chosen by either A or B c) chosen by exactly one of A and B	Let X_i = 1 if both choose item i, and 0 otherwise. Let Y_i = 1 if neither A nor B choose item i and 0 otherwise. Let W_i = 1 if exactly one of A and B choose item i and 0 otherwise. Let X = $\sum_{i=1}^{10}X_i$, Y = $\sum_{i=1}^{10}Y_i$, W = $\sum_{i=1}^{10}W_i$ a) E(X) = $\sum_{i=1}^{10}E(X_i)$ = 10$(3/10)^2$ = .9 b) E(Y) = $\sum_{i=1}^{10}E(Y_i)$ = 10$(7/10)^2$ = 4.9 c) Since X + Y + W = 10. E(W) = 10 - .9 - 4.9 = 4.2 OR $\sum_{i=1}^{10}E(W_i)$ = 10(2)(3/10)(7/10)
Cards are turned face up 1 at a time. If 1st card is ace, or 2nd a deuce, or 3rd a 3, or ... or 13th a King, or 14th an ace, and so on, we say that a match occurs.	Compute expected num of matches that occur E(number of matches) = E[$\sum_{i=1}^{52}I_i$], I_i = $\begin{cases}1, match\ on\ card\ i \\ 0, otherwise\end{cases}$ = 52E(I_i) = 52(1/13) = 4
Let X be a r.v. having finite expectation μ and variance σ^2, and let g be a twice differentiable fn. Shoe E(g(X)) ≈ g(μ) + $\frac{g''(\mu)}{2}\sigma^2$	g(X) = g(μ) + g'(μ)(X - μ) + g''(μ) $\frac{(X-\mu)^2}{2}$ + ... ≈ g(μ) + g'(μ)(X - μ) + g''(μ) $\frac{(X-\mu)^2}{2}$ E(g(X)) = E{g(μ) + g'(μ)(X - μ) + g''(μ) $\frac{(X-\mu)^2}{2}$} = g(μ) + g'(μ)E(X - μ) + g''(μ) $\frac{1}{2}$E$(X-\mu)^2$ = g(μ) + g'(μ){E(X) - μ} + $\frac{g''(\mu)}{2}\sigma^2$ = g(μ) + $\frac{g''(\mu)}{2}\sigma^2$
Let $X_1, X_2, ..., X_n$ be indep and identically distributed +ve r.v.. Find for k ≤ n, E$\left[\frac{\sum_{i=1}^k X_i}{\sum_{i=1}^n X_i}\right]$	1 = E$\left[\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i}\right]$ = E$\left[\frac{X_1+X_2+...+X_n}{\sum_{i=1}^n X_i}\right]$ = $\sum_{i=1}^n$ E$\left[\frac{X_i}{\sum_{i=1}^n X_i}\right]$ = nE$\left[\frac{X_1}{\sum_{i=1}^n X_i}\right]$. So E$\left[\frac{X_1}{\sum_{i=1}^n X_i}\right]$ = $\frac{1}{n}$, then E$\left[\frac{\sum_{i=1}^k X_i}{\sum_{i=1}^n X_i}\right]$ = $\frac{k}{n}$

If E(X) = 1 and Var(X) = 5, find $E[(2+X)^2]$ and Var(4 + 3X)	$E[(2+X)^2] = E[X^2 + 4X + 4] = E[X^2] + 4E(X) + 4 = \{Var(X) + [E(X)]^2\} + 4E(X) + 4 = 14$ Var(4 + 3X) = 9E(X) = 45
If 10 married couples are randomly seated at a round table, compute expected num and var of num of wives who are seated next to their husbands	Let $X_j = \begin{cases}1, if\ couple\ j\ are\ next\ to\ e.o.\\ 0, otherwise\end{cases}$. $E(\sum_{j=1}^{10} X_j) = \sum_{j=1}^{10} E(X_j) = 10[1*P(X_j = 1)] = 10(2/19) = 20/19$. (Since there are 2 ppl seated next to wife j, prob 1 of them is her husband is 2/19) $Var(\sum_{j=1}^{10} X_j) = \sum_{j=1}^{10} Var(X_j) + 2\binom{10}{2}Cov(X_i, X_j)$ = 10(2/19)(17/19) (var of Bernoulli) + $90[E(X_iX_j) - E(X_i)E(X_j)] = 340/361 + 90[P(X_i = 1, X_j = 1) - E(X_i)^2] = 340/361 + 90[P(X_i = 1)P(X_j = 1 \mid X_i = 1) - (2/19)^2] =$ 340/361 + 90[(2/19)(2/18) - 4/361] (couple 1 next to e.o., couple 2 need to be tgt and only have 18 seats left to choose from) = 360/361
Let X be num of 1's and Y be num of 2's that occur in n rolls of a fair die. Cov(X, Y)?	Let $X_i = \begin{cases}1, roll\ i\ is\ 1\\0, otherwise\end{cases}$, $Y_i = \begin{cases}1, roll\ i\ is\ 2\\0, otherwise\end{cases}$. $Cov(X_i, Y_j) = E[X_iY_j] - E(X_i)E(Y_j)$. If i = j: $X_iY_j = X_iY_i = 0$, since roll i is either 1, 2, or others. $Cov(X_i, Y_j) = -E(X_i)E(Y_j) = -P(X_i = 1)P(Y_j = 1) = -1/36$ If i ≠ j, $Cov(X_i, Y_j) = 0$. (indep since 2 diff rolls) $Cov(X, Y) = Cov(\sum_{i=1}^n X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^n \sum_{j=1}^n Cov(X_i, Y_j) = \sum_{i=1}^n Cov(X_i, Y_i) = -n/36$
Joint density fn of X and Y is f(x, y) = $\frac{1}{y}e^{-(y+\frac{x}{y})}$, x > 0, y > 0 Find E(X), E(Y) and show Cov(X, Y) = 1	$f_Y(y) = e^{-y}\int_0^\infty \frac{1}{y}e^{-x/y}\,dx = e^{-y}$, y > 0. Y ~ Exp(1). E(Y) = 1, Var(Y) = 1. $f_{X\mid Y}(x\mid y) = f(x,y)/f_Y(y) = \frac{1}{y}e^{-x/y}$. Then $X\mid Y=y$ ~ Exp(1/y). $E(X\mid Y=y) = y$. $E(X) = E(E(X\mid Y)) = E(Y) = 1$ $Cov(X, Y) = E(XY) - E(X)E(Y) = E[E(XY\mid Y)] - 1 = E[YE(X\mid Y)]$ (Since Y is a "constant" inside) $- 1 = E(Y^2) = 2 - 1 = 1$
Pond contains 100 fish, of which 30 are carps. If 20 fish are caught, what are the mean and var of num of carp among these 20.	Let X be num of carps caught. X ~ HGeo(20, 100, 30). E(X) = 20*30/100 = 6. Var(X) = (20*30)(100-30)(100-20)/(100^2*(100-1)) = 112/33
If X and Y are identically distributed, not nexessarily indep, show Cov(X + Y, X - Y) = 0	Cov(X+Y, X-Y) = Cov(X, X) - Cov(X, Y) + Cov(Y, X) - Cov(Y, Y) = Var(X) - Var(Y) = 0
Joint density of X and Y is given by f(x, y) = $\frac{e^{-x/y}e^{-y}}{y}$, 0 < x < ∞, 0 < y < ∞ Compute $E[X^2\mid Y = y]$	$f_{X\mid Y}(x\mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{e^{-x/y}e^{-y}}{y}}{\int_0^\infty \frac{e^{-x/y}e^{-y}}{y}dx} = \frac{e^{-x/y}}{y}$, 0 < x < ∞. $X\mid Y = y$ ~ Exp(1/y). $E(X\mid Y = y) = y$. $E[X^2\mid Y = y] = Var(X\mid Y = y) + [E(X\mid Y = y)^2] = y^2 + y^2 = 2y^2$
Joint density of X and Y is f(x, y) = $\frac{e^{-y}}{y}$, 0 < x < y, 0 < y < ∞. Compute $E[X^3\mid Y = y]$	$f_{X\mid Y}(x\mid y) = \frac{\frac{e^{-y}}{y}}{\int_0^\infty \frac{e^{-y}}{y}dx}$ = 1/y, 0 < x < y. $E[X^3\mid Y = y] = \int_0^y x^3(1/y)\,dx = y^3/4$
Expected num of accidents per week at an industrial plant is 5. Suppose num of workers injured in each accident are indep r.v. w common mean of 2.5. If num of workers injured in each accident is indep of num of accident occuring, compute expected num of workers injured in a week.	Let N be num of accidents, X_j be num of workers in accident j $E(X_1 + X_2 + ... + X_N) = E[E(X_1 + ... + X_N\mid N)]$ (since N is a r.v. as well) = E(2.5*N) = 2.5E(N) = 12.5
Type *i* light bulbs fn for a random amt of time w mean μ_i and sd σ_i, i = 1,2. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb w prob p, type 2 bulb w prob 1-p. Let X denote lifetime of bulb. Find E(X), Var(X)	$E(X) = E(X\mid type\ 1)p + E(X\mid type\ 2)(1-p) = p\boldsymbol{\mu}_1 + (1-p)\boldsymbol{\mu}_2$. Let I be r.v. denoting type of light bulb $Var(X) = E[Var(X\mid I)] + Var[E(X\mid I)] = E[\sigma_I^2] + Var(\boldsymbol{\mu}_I) = p\sigma_1^2 + (1-p)\sigma_2^2 + \{E(\boldsymbol{\mu}_I^2) - [E(\boldsymbol{\mu}_I)]^2\} = p\sigma_1^2 + (1-p)\sigma_2^2 + \{p\boldsymbol{\mu}_1^2 + (1-p)\boldsymbol{\mu}_2^2 - [p\boldsymbol{\mu}_1 + (1-p)\boldsymbol{\mu}_2]^2\}$
Num of accidents a person has in a given year is a Poisson r.v. w mean λ. Suppose value of λ is 2 for 60% of pop and 3 for other 40%. If person is chosen at random, prob that he will have 0 accidents, exactly 3 accidents in a year. What is the conditional prob that he will have 3 accidents in a given year, given he has no accidents in the previous year?	P(0 accidents) = $.6e^{-2} + .4e^{-3}$ P(3 accidents) = $.6e^{-2}(2^3/3!) + .4e^{-3}(3^3/3!)$ $P(3\ accidents\mid 0) = \frac{P(3,0)}{P(0)} = \frac{P(3\ accidents)}{P(0\ accidents)}$ (since accidents in previous year don't affect curr year)
Mgf of X is $M_X(t) = exp(2e^t - 2)$ and Y is $M_Y(t) = \left(\frac{3}{4}e^t + \frac{1}{4}\right)^{10}$. If X and Y are indep, what are P(X + Y = 2), P(XY = 0), E(XY)	X is poisson w $\boldsymbol{\lambda}$ = 2. Y is Binomial w param (10, 3/4) P(X + Y = 2) = P(X=0)P(Y=2) + P(X = 1)P(Y=1) + P(X=2)P(Y=0) P(XY=0) = P(X=0) + P(Y=0) - P(X=0, Y=0) E(XY) = E(X)E(Y) = 2*10*3/4 = 15
Joint density of X and Y is f(x,y) = $\frac{1}{\sqrt{2\pi}}e^{-y}e^{-\frac{(x-y)^2}{2}}$, 0 < y < ∞, -∞ < x < ∞. Compute joint mgf of X and Y. Compute individual mgf	Note that Y is exp w rate 1, and given Y, X is normal w var 1 $E[e^{tX+sY}] = E[e^{tX+sY}\mid Y] = e^{sY}E[e^{tX}\mid Y] = e^{sY}e^{Yt+t^2/2}$. $E[e^{tX+sY}] = E\{E[e^{tX+sY}\mid Y]\} = E\{e^{sY}e^{Yt+t^2/2}\} = e^{t^2/2}E[e^{(s+t)Y}] = e^{t^2/2}\frac{1}{1-(s+t)}$, s + t < 1 $E(e^{tX}) = e^{t^2/2}\frac{1}{1-t}$, t < 1 (let s = 0). $E(e^{sY}) = \frac{1}{1-s}$, s < 1

Markov's Inequality. If X is a r.v. that takes only nonnegative values, then for any a > 0, P(X ≥ a) ≤ $\frac{E(X)}{a}$	Proof. For. a > 0, let I = $\begin{cases}1, if\ X \geq a\\0, otherwise\end{cases}$. Note I ≤ X/a. E(I) ≤ E(X)/a E(I) = 1P(X ≥ a) + 0P(X < a) = P(X ≥ a). So P(X ≥ a) ≤ E(X)/a
Chebyshev's Inequality. If X is a r.v. w finite mean μ and var σ^2, then for any value of k > 0, $P(\lvert X - \mu\rvert \geq k) \leq \frac{\sigma^2}{k^2}$	$(X - \mu)^2 \geq 0$. $P[(X - \mu)^2 \geq k^2] \leq \frac{E[(X-\mu)^2]}{k^2}$ (from Markov's Inequality) $P(\lvert X- \mu\rvert \geq k) \leq \frac{\sigma^2}{k^2}$ (since $E[(X-\mu)^2] = var(X) = \sigma^2$)
Let X have pdf f(x) = $\begin{cases}\frac{1}{2\sqrt{3}}, -\sqrt{3} < x < \sqrt{3}\\ 0, otherwise\end{cases}$ Find $P(\lvert X\rvert \geq 3/2)$ exactly and approximately using Chebyshev's inequality	E(X) = 0 (since symmetric). Var(X) = $\frac{(\beta-\alpha)^2}{12} = \frac{(\sqrt{3}-(-\sqrt{3}))^2}{12} = 1$ $P(\lvert X\rvert > 3/2) = 1 - \int_{-3/2}^{3/2}\frac{1}{2\sqrt{3}}dx = 0.134$ $P(\lvert X\rvert > 3/2) = P(\lvert X-0\rvert > 3/2) \leq \frac{1}{(3/2)^2} = 0.444$ (Chebyshev's)
If Var(X) = 0, then P(X = E[X]) = 1	Proof. Let μ = E(X). Using Chebyshev's inequality, for any n ≥ 1, $P(\lvert X- \mu\rvert \geq \frac{1}{n}) \leq \frac{Var(X)}{(1/n)^2} = 0$ So $P(\lvert X- \mu\rvert \geq \frac{1}{n}) = 0$. $0 = \lim_{n\to\infty} P(\lvert X - \mu\rvert \geq \frac{1}{n}) = P(\lim_{n\to\infty}(\lvert X - \mu\rvert \geq \frac{1}{n})) = P(X \neq \mu)$. So $P(X = \mu) = 1$
Weak law of large nums. Let $X_1, X_2,...$ be a seq of indep and identically distributed r.v. each having finite mean $E[X_i] = \mu$. Then for any $\varepsilon > 0$, $P\left\{\left\lvert\frac{X_1+\cdots+X_n}{n} - \mu\right\rvert \geq \varepsilon\right\} \to 0$ as n → ∞	Proof. $E\left[\frac{X_1+\cdots+X_n}{n}\right] = \mu$. $Var(\frac{X_1+\cdots+X_n}{n}) = \frac{\sigma^2}{n}$. Using Chebyshev's inequality, $P\left\{\left\lvert\frac{X_1+\cdots+X_n}{n} - \mu\right\rvert \geq \varepsilon\right\} \leq \frac{\sigma^2}{n\varepsilon^2} \to 0$ as n → ∞
Let $X_1, X_2, ...$ be a seq of r.v. s.t. $X_n \sim N(\mu + \frac{1}{n}, \sigma^2)$. If $X_n \to X$, what is the dist of X? Note $X \sim N(\mu, \sigma^2)$. $M(t) = e^{t^2\sigma^2/2 + t\mu}$	$X_n \sim N(\mu + \frac{1}{n}, \sigma^2)$. $M_{X_n}(t) = e^{(\mu+\frac{1}{n})t + t^2\sigma^2/2} \to e^{\mu t + t^2\sigma^2/2}$ = mgf of $N(\mu, \sigma^2)$. So $X \sim N(\mu, \sigma^2)$.

CLT. Let $X_1, X_2,\ldots$ be a seq of indep and identically distributed r.v. each having mean μ and var σ^2. Then $\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ tends to standard normal as n → ∞. Note (Var(X) = E(X^2) - [E(X)]^2. 1 = E(X^2) - 0.)	Proof. Assume $\mu = 0$, $\sigma^2 = 1$. $\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}} = \frac{X_1+\cdots+X_n}{\sqrt{n}} = \frac{X_1}{\sqrt{n}} + \ldots + \frac{X_n}{\sqrt{n}}$ M = $M_{\frac{X_1}{\sqrt{n}}+\ldots+\frac{X_n}{\sqrt{n}}}(t) = M_{\frac{X_1}{\sqrt{n}}}(t)\ldots M_{\frac{X_n}{\sqrt{n}}}(t)$ (since X_i are indep) $M_{\frac{X_i}{\sqrt{n}}}(t)$ = E$\left[e^{t\frac{X_i}{\sqrt{n}}}\right]$ = E$\left[e^{\frac{t}{\sqrt{n}}X_i}\right]$ = $M_{X_i}(\frac{t}{\sqrt{n}}) = M_X(\frac{t}{\sqrt{n}})$. So M = $\left[M_X(\frac{t}{\sqrt{n}})\right]^n$ Let L(t) = log[$M_X(t)$] where $M_X(t)$ = E[e^{tX}]. Then L(0) = log[$M_X(0)$] = log 1 = 0 L'(t) = $\frac{M_X'(t)}{M_X(t)}$. L'(0) = $\frac{M_X'(0)}{M_X(0)} = \frac{\mu}{1}$ = 0. L''(t) = $\frac{M_X(t)M_X''(t)-[M_X'(t)]^2}{[M_X(t)]^2}$. L''(0) = $\frac{1E(X^2)-[\mu]^2}{[1]^2}$ = E(X^2) = 1. So L($\frac{t}{\sqrt{n}}$) = log$\left[M_X(\frac{t}{\sqrt{n}})\right]$. $e^{\mathrm{L}(\frac{t}{\sqrt{n}})} = M_X(\frac{t}{\sqrt{n}})$. M = $\left[M_X(\frac{t}{\sqrt{n}})\right]^n = \left[e^{\mathrm{L}(\frac{t}{\sqrt{n}})}\right]^n = e^{\mathrm{nL}(\frac{t}{\sqrt{n}})}$ Need to show $\lim_{n\to\infty} M = \lim_{n\to\infty} e^{\mathrm{nL}(\frac{t}{\sqrt{n}})} = e^{t^2/2}$. Same as showing $\lim_{n\to\infty} \mathrm{nL}(\frac{t}{\sqrt{n}})$ = $t^2/2$ $\lim_{n\to\infty} \mathrm{nL}(\frac{t}{\sqrt{n}}) = \lim_{n\to\infty} \frac{\mathrm{L}(\frac{t}{\sqrt{n}})}{1/n} = \lim_{n\to\infty} \frac{-\mathrm{L}'(\frac{t}{\sqrt{n}})n^{-3/2}t}{-2n^{-2}}$ (L'Hopital's rule) = $\lim_{n\to\infty} \frac{\mathrm{L}'(\frac{t}{\sqrt{n}})t}{2n^{-1/2}} = \lim_{n\to\infty} \frac{-\mathrm{L}''(\frac{t}{\sqrt{n}})n^{-3/2}t^2}{-2n^{-3/2}}$ = $\lim_{n\to\infty} L''(\frac{t}{\sqrt{n}})\frac{t^2}{2}$ = $t^2/2$
Suppose a seq of indep trials is performed. Let E be a fixed event and prob occur is P(E). Let $X_i = \begin{cases} 1, if\ E\ occurs\ on\ ith\ trial \\ 0, if\ E\ don't\ occur\ on\ ith\ trial \end{cases}$. Find $\lim_{n\to\infty} \frac{X_1+\cdots+X_n}{n}$	Using strong law of large nums, $\frac{X_1+\cdots+X_n}{n}$ → E(X_1) with prob 1 E(X_1) = 1P(E) + 0P(E^C) = P(E) In other words, $\frac{num\ of\ times\ E\ occurs}{total\ num\ of\ trials}$ → P(E) with prob 1
Chernoff bounds. P(X ≥ a) ≤ e^{-ta}M(t) for all t > 0.	Proof. P(X ≥ a) = P(tX ≥ ta) for t > 0 = P($e^{tX} \geq e^{ta}$) ≤ E(e^{tX})/e^{ta} (Markov's inequality) = e^{-ta}M(t) for all t > 0.
Z ~ N(0,1), find Chernoff bound for Z	P(Z ≥ a) ≤ $e^{-ta}M_Z(t)$ for t > 0 = $e^{-ta}e^{t^2/2} = e^{t^2/2-ta}$ for t > 0 h(t) = $t^2/2$ - ta. h'(t) = t-a. h'(t) = 0, then t = a. h''(t) = 1 > 0 (min value). So P(Z ≥ a) ≤ $e^{-a^2/2}$ (tightest bound)
Jensen's inequality If f(x) is a convex fn, then E[f(X)] ≥ f(E(X)), if E(X) exists and is finite	Proof. Let μ = E(X). Taylor's series expansion: f(x) = f(μ) + f'(μ)(x-μ) + $\frac{f''(\xi)(x-\mu)^2}{2}$ where $\xi \in$ (x, μ) $\frac{f''(\xi)(x-\mu)^2}{2}$ ≥ 0 since convex fn. So f(x) ≥ f(μ) + f'(μ)(x-μ). f(X) ≥ f(μ) + f'(μ)(X-μ). E(f(X)) ≥ E{f(μ) + f'(μ)(X-μ)} = f(μ) + f'(μ)(E(X)-μ) = f(E(X)) + f'(μ)(E(X)-E(X)) = f(E(X))
Let X be a positive r.v.. Show E(1/X) ≥ 1/E(X)	f(x) = 1/x. f'(x) = -x^{-1}. f''(x) = x^{-2} > 0 for all x > 0. f is convex fn. By Jensen's inequality, E(1/X) ≥ 1/E(X)
Let X be a r.v. w P(X ≤ 0) = 0. i.e. X is +ve. Show P(X ≥ 2μ) ≤ 1/2 where μ = E(X)	$\mu = \int_0^\infty xf(x)\,dx \geq \int_{2\mu}^\infty xf(x)\,dx \geq \int_{2\mu}^\infty 2\mu f(x)\,dx$ (since smallest value x can take is 2μ) = 2μP(X ≥ 2μ) P(X ≥ 2μ) ≤ 1/2
Extra	
Let $X_1, \ldots, X_{20}$ be indep Poisson r.v. w mean 1. Use Markov inequality to obtain a bound on P($\sum_{i=1}^{20} X_i > 15$). Use clt to approximate P($\sum_{i=1}^{20} X_i > 15$).	P($\sum_{i=1}^{20} X_i > 15$) ≤ E($\sum_{i=1}^{\mathbf{20}} \boldsymbol{X_i}$)/15 = 20/15 Sum ~ N(20, 20). P($\sum_{i=1}^{20} X_i > 15$) = P($\sum_{i=1}^{20} X_i \geq 15.5$) = .8428
A die is continually rolled until total sum of all rolls exceeds 300. Prob that at least 80 rolls are necessary?	If X_i is outcome on i^{th} roll, then E(X_i) = 7/2, Var(X_i) = 35/12. P(at least 80 rolls are needed) = P(sum ≤ 300) = P($\sum_{i=1}^{79} X_i \leq 300.5$) = 0.9430 (normal approx)
If X is a gamma r.v. w param (n, 1), how large does n need to be s.t. P($\left\|\frac{X}{n} - 1\right\| > .01$) < .01	Gamma(n,1) = sum of n indep exponential variables w rate 1, thus X has mean n, var n. P($\left\|\frac{X}{n} - 1\right\| > .01$) = P($\left\|\frac{X-n}{\sqrt{n}}\right\| > .01\sqrt{n}$) = 2P(Z > .01$\sqrt{n}$) 2P(Z > .01$\sqrt{n}$) < .01. P(Z > .01$\sqrt{n}$) < .005. Using normal table, .01$\sqrt{n}$ = 2.58. n = 258^2