NAME: Answers

Math 461 Fall 2006 — Final Exam

Total points: **200**. Do all questions. **Explain** all answers. You may use any formulas from class for the density, expectation or variance of standard random variables. No notes, books, or electronic devices.

Useful formulas.

Geometric series: $$\sum_{x=0}^{\infty} r^x = \frac{1}{1-r}, \qquad -1 < r < 1,$$

$$\int_{-c}^{c} \frac{1}{\sqrt{1-x^2}}\,dx = 2\arcsin(c)$$

$$\int_{-c}^{c} \sqrt{1-x^2}\,dx = c\sqrt{1-c^2} + \arcsin(c)$$

---

1. [15 points] An airline owns 3 types of plane: Airbus, Boeing and Embraer, with 30 planes of each type.

Find the probability that a random group of 10 planes will be void in at least one type. (*Hint.* There are at least two different solution methods.)

(i) $P(\text{void in} \geq 1 \text{ type})$

$= P(A \cup B \cup E)$ where A: void in Airbus, B: " " Boeing, E: " " Embraer

$= P(A) + P(B) + P(E) - [P(AB) + P(BE) + P(EA)] + P(ABE)$ by Inclusion-Exclusion

$= 3P(A) - 3P(AB) + 0$ ← since ABE is impossible

$= 3\frac{\binom{60}{10}}{\binom{90}{10}} - 3\frac{\binom{30}{10}}{\binom{90}{10}}$ ← (void in A and B means only 30 planes to choose from)

← (void in A means only 60 planes to choose from)

(ii) $1 - P(\text{void in 0 types}) = 1 - \sum_{\substack{j,k,l \geq 1: \\ j+k+l=10}} \binom{30}{j}\binom{30}{k}\binom{30}{l} \Big/ \binom{90}{10}$

2. [25=15+10 points] Memory chips are defective with probability $p = .001$, independently of one another.

(a) Show the probability of a batch of n chips containing at most one defective is approximately $e^{-np}(1+np)$. Explain what kind of random variables you are using, and why.

$X \sim \text{Binomial}(n, p)$, but p is small and n is large,

so $X \sim \text{Poisson}(np)$ is a good approximation.

$\therefore P(X \leq 1) = P(X=0) + P(X=1)$, and we use the Poisson:

$$= e^{-\lambda}\frac{\lambda^0}{0!} + e^{-\lambda}\frac{\lambda^1}{1!} = e^{-np}(1+np).$$

(b) A batch of 1000 chips contains more than one defective. Show the probability it contains exactly two defectives is approximately $1/(2(e-2))$.

$P(2 \text{ defective} \mid >1 \text{ defective})$

$$= \frac{P(2 \text{ defective and } >1 \text{ defective})}{P(>1 \text{ defective})}$$

$P(>1 \text{ defective}) = 1 - P(X \leq 1)$

$$= \frac{e^{-\lambda}\lambda^2/2!}{1 - e^{-\lambda}(1+\lambda)} = \frac{e^{-1}/2}{1-e^{-1}(1+1)} = \frac{1}{2(e-2)}$$

since $np = 1000p = 1$

(c) [*Extra credit*, 5 points.] The manufacturer requires that at least 98% of batches contain at most one defective. What is the largest allowable batch size?

$$.98 \leq P(X \leq 1) = e^{-np}(1+np) \quad \text{by part (a)}$$

$$= \left(1 - np + \frac{(np)^2}{2!} + \dots\right)(1+np)$$

$$= 1 - \frac{(np)^2}{2} + \dots$$

$\therefore \frac{(np)^2}{2} \leq .02$, so $np \leq .2$

ie $\boxed{n \leq 200}$

3. [25=14+6+5 points]

(a) Suppose X is a continuous random variable with density $f(x)$. Find the density $g(y)$ of the random variable $Y = 4X$.

$$P(a<Y<b) = \int_a^b g(y)\,dy$$

But $$P(a<Y<b) = P\left(\frac{a}{4} < X < \frac{b}{4}\right)$$

$$= \int_{a/4}^{b/4} f(x)\,dx \,, \quad \text{now let } x = \frac{y}{4}$$

$$= \int_a^b f(y/4)\,\frac{dy}{4}$$

$$\therefore \quad g(y) = f\left(\frac{y}{4}\right)\cdot\frac{1}{4}.$$

(b) Assume $X \sim \text{Gamma}(\alpha, \lambda)$. State the density $f(x)$. Then use part (a) to get an explicit formula for the density $g(y)$ of $Y = 4X$.

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)}\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}, & \text{if } x>0 \\ 0, & \text{if } x \le 0 \end{cases}$$

So $$g(y) = \frac{1}{4} f\left(\frac{y}{4}\right)$$

$$= \begin{cases} \frac{1}{\Gamma(\alpha)}\frac{\lambda}{4} e^{-\lambda y/4}\left(\lambda \frac{y}{4}\right)^{\alpha-1}, & \text{if } y>0 \\ 0, & \text{otherwise} \end{cases}$$

(c) Give another derivation of the density $g(y)$, in part (b), by considering the intuitive meaning of Gamma random variables.

$X \sim \text{Gamma}(\alpha, \lambda)$ means $X =$ waiting time for αth event, for events that happen at average rate λ.

So $Y = 4X$ means $Y =$ waiting time for αth event, for events that happen at rate $\frac{\lambda}{4}$. (Note that if events happen less often, then you have to wait longer for them.)

4. [20 points] Show that the expected distance between two independently chosen points in the interval $0 < x < 1$ (using the uniform probability density) is 1/3.

Let X = first point ~ Uniform(0,1)

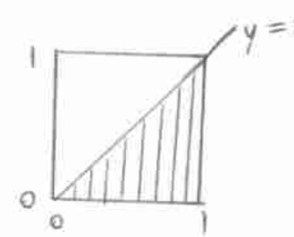
Y = second point ~ Uniform(0,1)

be independent. Then their joint density on the unit square is $1 \cdot 1 = 1$.

And E[distance from X to Y]

$$= E[\,|X-Y|\,]$$

$$= \int_0^1 \int_0^1 |x-y| \, dx \, dy$$



$$= 2 \int_0^1 \int_0^x (x-y) \, dy \, dx$$

by symmetry (considering the regions where $y<x$ and $y>x$ in turn)

$$= 2 \int_0^1 \left[xy - \tfrac{1}{2}y^2 \right]_{y=0}^{y=x} dx$$

$$= 2 \int_0^1 \tfrac{1}{2}x^2 \, dx$$

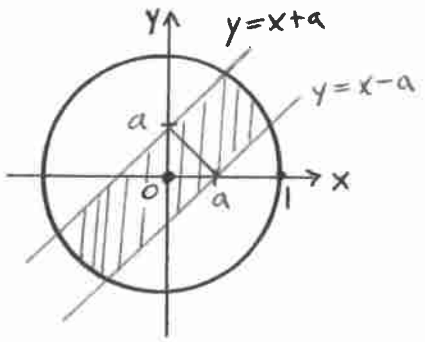
$$= \tfrac{1}{3}x^3 \Big|_0^1$$

$$= \tfrac{1}{3}.$$

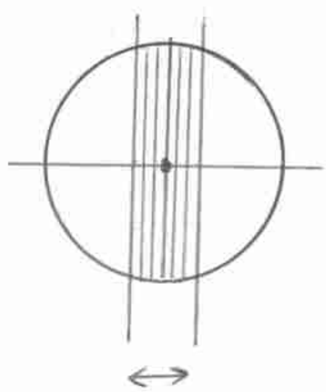
5. [20 points] Choose a point (X, Y) randomly in the disk of radius 1 centered at the origin, using the uniform density on the disk.

Evaluate $P(|Y - X| < a)$, for $0 < a < \sqrt{2}$.

Hint. Rotate.



rotate 45 degrees



$$P(|Y-X|<a) = P(X-a<Y<X+a)$$

$$= \frac{1}{\pi}\cdot(\text{shaded area}),$$ since we use the uniform density on the disk

$$= \frac{1}{\pi}\cdot(\text{shaded area in rotated diagram})$$

$$= \frac{1}{\pi}\cdot\int_{-a/\sqrt{2}}^{a/\sqrt{2}}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy\,dx$$

$$= \frac{2}{\pi}\int_{-a/\sqrt{2}}^{a/\sqrt{2}} \sqrt{1-x^2}\,dx$$

$$= \frac{2}{\pi}\left(\frac{a}{\sqrt{2}}\sqrt{1-\left(\frac{a}{\sqrt{2}}\right)^2} + \arcsin\left(\frac{a}{\sqrt{2}}\right)\right)$$

by a Useful Formula on front of exam.

width $= \sqrt{a^2+a^2}$ by Pythagoras

$= \sqrt{2}\,a$

$= 2\cdot\frac{a}{\sqrt{2}}$

Use this in the limits of integration.

6. [10 points] **Do part (a) or part (b), but not both.**

(a) Assume $X \sim$ Exponential(2) and $Y \sim$ Exponential(1) are independent. Show $P(Y \le 3X) = 3/5$.

$P(Y \le 3X)$

$= \int_0^\infty \int_0^{3x} 2e^{-2x} \cdot 1e^{-1y}\, dy\, dx$ (joint density = product of individual densities, by independence)

$= \int_0^\infty 2e^{-2x}(1-e^{-3x})\, dx$

$= \int_0^\infty (2e^{-2x} - 2e^{-5x})\, dx$

$= 1 - \frac{2}{5} = \frac{3}{5}$

(b) Assume $X \sim$ Normal(20, 9) and $Y \sim$ Normal(50, 19) are independent. Show $P(Y \le 3X) \simeq .8413$.

$-3X \sim \text{Normal}(-3 \cdot 20, (-3)^2 \cdot 9)$

$= \text{Normal}(-60, 81)$

$Y \sim \text{Normal}(50, 19)$

$\therefore Y - 3X \sim \text{Normal}(-10, 100)$ by independence

So $P(Y-3X \le 0) = P\left(\frac{Y-3X-(-10)}{\sqrt{100}} \le \frac{0-(-10)}{\sqrt{100}} \right)$ (subtract the mean and divide by the standard deviation)

$= P(Z \le 1)$

$= \Phi(1)$

$= .8413.$

7. [25=8+12+5 points] A pond contains n fish, of which 30 are carp. (Assume $n \geq 30$.) Then 20 fish are caught (randomly selected). Let X be the number of carp among the fish caught. Write $e(n) := E[X]$ and $v(n) := \text{Var}(X)$ for the expectation and variance of X (these obviously depend on n).

(a) Use indicator random variables to show $e(n) = 30 \cdot 20/n$.

Let $X_i = \begin{cases} 1 & \text{if carp } i \text{ is caught} \\ 0 & \text{otherwise} \end{cases}$, for $i = 1, \ldots, 30$.

$$X = X_1 + \cdots + X_{30} = \sum_{i=1}^{30} X_i$$

$$e(n) = E[X] = \sum_{i=1}^{30} E[X_i] = \sum_{i=1}^{30} P(X_i = 1) = \sum_{i=1}^{30} \frac{20}{n}$$

since 20 fish are caught out of n total fish.

$\therefore e(n) = 30 \cdot 20/n$.

(b) Find a formula for $v(n)$. (You need not simplify your formula.)

$$v(n) = \text{Var}(X)$$
$$= \text{Var}(X_1 + \cdots + X_{30})$$
$$= \sum_{i=1}^{30} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$
$$= 30 \,\text{Var}(X_1) + 30 \cdot 29 \,\text{Cov}(X_1, X_2) \quad \text{by symmetry}$$

$$v(n) = 30 \cdot \frac{20}{n}\left(1 - \frac{20}{n}\right) + 30 \cdot 29 \left(\frac{20}{n} \cdot \frac{19}{n-1} - \left(\frac{20}{n}\right)^2 \right)$$

since $\text{Var}(X_1) = p(1-p) = \frac{20}{n}\left(1 - \frac{20}{n}\right)$ and $\text{Cov}(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$, with

$$E[X_1 X_2] = P(X_1 X_2 = 1) = P(X_2 = 1, X_1 = 1) = P(X_2 = 1 \mid X_1 = 1) \cdot P(X_1 = 1) = \frac{19}{n-1} \cdot \frac{20}{n}.$$

(c) Now assume the total number of fish is not a fixed number n, but is a random variable N. (Assume $N \geq 30$, always.)

Write down formulas for the expectation and variance of the number of carp caught. Your answers should use the functions $e(\cdot)$ and $v(\cdot)$ somehow.

$$E[X] = \sum_{n=30}^{\infty} E[X \mid N = n] P(N = n) \quad \text{by conditioning on } N$$
$$= E[e(N)], \text{ since } E[X \mid N = n] = e(n)$$

And

$$\text{Var}(X) = E[\text{Var}(X \mid N)] + \text{Var}(E[X \mid N])$$
$$= E[v(N)] + \text{Var}(e(N)).$$

8. [20 points] Ten hunters are waiting for ducks to fly by. When a flock of ducks flies overhead, the hunters fire at the same time. Each hunter chooses his target at random, independently of the others. Assume that the number of ducks in a flock is a Poisson random variable with mean 6.

If each hunter independently hits his target with probability .6, then compute the expected number of ducks that are hit.

Let $X = \#\text{ducks} \sim \text{Poisson}(6)$.

Then

$E[\text{ducks hit}]$

$= \sum_{x=0}^{\infty} E[\text{ducks hit} \mid X=x]\, P(X=x)$

$= \sum_{x=0}^{\infty} E[\sum_{i=0}^{x} Y_i \mid X=x]\, P(X=x)$ where $Y_i = \begin{cases} 1 & \text{if duck } i \text{ hit} \\ 0 & \text{otherwise} \end{cases}$

$= \sum_{x=0}^{\infty} x\, E[Y_1 \mid X=x]\, P(X=x)$ by linearity, and symmetry,

$= \sum_{x=0}^{\infty} x\, P(Y_1 = 1 \mid X=x)\, P(X=x)$

$= \sum_{x=0}^{\infty} x \{1 - P(Y_1 = 0 \mid X=x)\}\, P(X=x)$

$= \sum_{x=0}^{\infty} x \{1 - P(\text{duck 1 not hit} \mid x \text{ ducks})\}\, P(X=x)$

$= \sum_{x=0}^{\infty} x \{1 - (1 - \frac{.6}{x})^{10}\}\, P(X=x)$

$P(X=x) = e^{-6}\frac{6^x}{x!}$, Poisson

duck 1 is not hit
$\Leftrightarrow$ all 10 hunters do not hit duck 1, and
$1 - \frac{.6}{x}$ represents the probability
of an individual hunter not hitting duck 1
(since $\frac{.6}{x}$ is the prob. of hitting duck 1)

9. [20=8+12 points] **Do #9 or #10 but not both.** From past experience, a professor knows that the test score of a student taking her final examination is a random variable with mean 75.

(a) Show that the probability a student's test score will exceed 85 is less than 75/85. (State the name of any inequality you use.)

$$P(X \geq 85) \leq \frac{E[X]}{85} = \frac{75}{85}$$

by the Markov Ineq., noting X is nonnegative and $\mu = 75$

(b) Suppose in addition that the professor knows the variance of a student's score is equal to 25. What can be said about the probability that the student will score between 65 and 85? (State the name of any inequality you use.)

$\mu = 75$, $\sigma^2 = 25$, $\sigma = 5$

$$P(|X-75| > 10)$$
$$= P(|X-\mu| > 2\sigma)$$
$$\leq \frac{1}{2^2} = \frac{1}{4}$$ by Chebyshev's inequality

So $$P(|X-75| \leq 10) = 1 - P(|X-75| > 10) \geq 1 - \frac{1}{4} = \frac{3}{4}.$$

10. [20=8+12 points] **Do #9 or #10 but not both.** Forty-eight numbers are chosen independently, according to the uniform distribution on the interval $(-0.5, 0.5)$. Write S for the sum of the numbers.

(a) Find $E[S]$ and $\text{Var}(S)$.

$X_i \sim \text{Uniform}(-.5, .5)$, $\quad E[X_i] = 0$, $\text{Var}(X_i) = \frac{1}{12}$

$S = X_1 + \cdots + X_{48}$

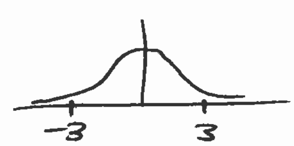
So
$$E[S] = E[X_1] + \cdots + E[X_{48}] \quad \text{by linearity}$$
$$= 0 + \cdots + 0$$
$$= 0$$

and
$$\text{Var}(S) = \text{Var}(X_1) + \cdots + \text{Var}(X_{48}) \quad \text{by independence}$$
$$= \frac{1}{12} + \cdots + \frac{1}{12}$$
$$= 48 \cdot \frac{1}{12}$$
$$= 4$$

(b) Show $P(|S| > 3) \simeq .1336$.

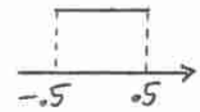
$S \approx \text{Normal}(0, 4)$ by the Central Limit Theorem (using independence of $X_1, \ldots, X_{48}$)

Hence
$$P(|S| > 3)$$
$$= P(S > 3) + P(S < -3)$$



$$= 2P(S > 3)$$
$$= 2(1 - P(S < 3))$$
$$= 2\left(1 - P\left(\frac{S-0}{\sqrt{4}} < \frac{3-0}{\sqrt{4}}\right)\right)$$
$$\simeq 2\left(1 - \Phi\left(\frac{3}{\sqrt{4}}\right)\right) \quad \text{since} \quad \frac{S-0}{\sqrt{4}} \simeq \text{Normal}(0,1)$$
$$= 2(1 - \Phi(3/2))$$
$$= 2(1 - .9332)$$
$$= .1336$$

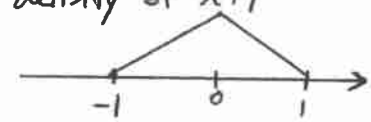
density of X



11. [20=15+5 points] Let X and Y be independent, Uniform$(-0.5, 0.5)$ random variables.

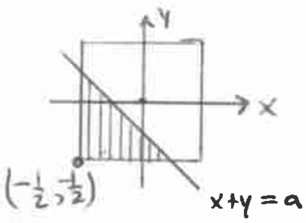
(a) Find the density of $X+Y$. You should check it has graph:

density of X+Y



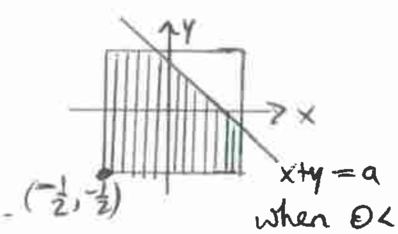
First we find the distribution function:

$$P(X+Y \le a) = \iint_{\{\text{shaded region}\}} \underbrace{\frac{1}{.5-(-.5)} \frac{1}{.5-(-.5)}}_{\text{joint density of } X \text{ and } Y} dx\,dy$$



$= $ area of shaded region

$$= \begin{cases} \frac{1}{2}(a+1)^2, & \text{when } -1<a<0 \\ 1-\frac{1}{2}(1-a)^2, & \text{when } 0<a<1 \end{cases}$$



We differentiate the distribution fn. with respect to a, to get the density:

$$\begin{cases} a+1, & \text{if } -1<a<0 \\ 1-a, & \text{if } 0<a<1 \end{cases}$$

which matches the graph.

(b) Squint at the graph in part (a), and then write a short explanation of how it relates to the Central Limit Theorem.

The Central Limit Theorem says that a sum of independent, identically distributed random variables will have approximately a normal distribution.

In part (a) we are adding two independent, identically distributed random variables, and indeed if you squint at the triangular density graph for X+Y, you can see (roughly) the shape of a normal density.

12. [*Extra credit,* 10 points] $X \sim$ Geometric(p) has moment generating function

$$M(t) = \frac{pe^t}{1-(1-p)e^t}.$$

Find the moment generating function of $Y \sim$ Neg. Binomial(r, p). Explain.

Write $Y = X_1 + \cdots + X_r$

where $X_1, \ldots, X_r$ are independent

Geometric r.v. Then

$$M_Y = M_{X_1} \cdots M_{X_r} = M^r.$$

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF x

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998