Math 480 Exam 1 is Wed. Sept. 23. **You are allowed 7 sheets of notes and a calculator.** The exam covers HW1-4, and Q1-4. Numbers refer to types of problems on exam. In this class $\log(t) = \ln(t) = \log_e(t)$ while $\exp(t) = e^t$. Note that $F(F^{-1}(u)) = F(x_u) = u$ where $0 < u < 1$. Hence $x_u = F^{-1}(u)$ is the 100 uth percentile of X.

0) Get familiar with the following distributions. For continuous distributions, assume formulas are given on the support.

a) beta(α, β): The **support** is [0,1]. The pdf $f(x) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ where $\alpha > 0$ and $\beta > 0$. $E(X) = \dfrac{\alpha}{\alpha+\beta}$. $V(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

b) chi-square(k) = gamma$(\alpha = k/2, \lambda = 1/2)$

$$f(x) = \frac{x^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}$$

where $x > 0$ and k is a positive integer.
$E(X) = k, \quad V(X) = 2k$. The mgf

$$\phi(t) = \left(\frac{1}{1-2t}\right)^{k/2} = (1-2t)^{-k/2}$$

for $t < 1/2$.

c) Exponential(λ)= Gamma$(\alpha = 1, \lambda)$: $f(x) = \lambda e^{-\lambda x}$ where x, $\lambda > 0$.
$F(x) = 1 - e^{-\lambda x}, \quad E(X) = 1/\lambda, \quad V(X) = 1/\lambda^2,$
$E(X^k) = \Gamma(k+1)/\lambda^k$ for $k > -1$. If k is a positive integer, $E(X^k) = k!/\lambda^k$.
$\phi(t) = (1 - t/\lambda)^{-1} = \dfrac{\lambda}{\lambda - t}, \ t < \lambda. \quad F^{-1}(u) = -\ln(1-u)/\lambda$ for $0 < u < 1$.

d) Gamma(α, λ): $f(x) = \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ where α, λ, and x are positive.
$E(X) = \alpha/\lambda, \quad V(X) = \alpha/\lambda^2, \quad E(X^k) = \dfrac{\Gamma(\alpha+k)}{\lambda^k \Gamma(\alpha)}$ for $k > -\alpha$.
$\phi(t) = (1 - t/\lambda)^{-\alpha} = \left(\dfrac{\lambda}{\lambda - t}\right)^\alpha$ for $t < \lambda$.

e) normal $N(\mu, \sigma^2)$: $E(X) = \mu$, $V(X) = \sigma^2$. The **support** is $(-\infty, \infty)$. If $Z \sim N(0,1)$, then the cdf of Z is $\Phi(x)$. If $X \sim N(\mu, \sigma^2)$, then the cdf of X is $F(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$, and the pdf $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[\dfrac{-1}{2\sigma^2}(x-\mu)^2\right]$. Then $F^{-1}(u) = \mu + \sigma z_u$ for $0 < u < 1$ where $\Phi(z_u) = u$. Here $\sigma > 0$ and μ is real. $\phi(t) = \exp(t\mu + t^2\sigma^2/2)$.

f) Uniform(α, β). This distribution has **support** on $\alpha \le x \le \beta$, $f(x) = \dfrac{1}{\beta - \alpha}$, $F(x) = (x-\alpha)/(\beta-\alpha)$, $E(X) = (\alpha+\beta)/2$, $V(X) = (\beta-\alpha)^2/12$, $\phi(t) = \dfrac{e^{t\beta} - e^{t\alpha}}{(\beta-\alpha)t}$, for $t \ne 0$ with $\phi(0) = 1$. $F^{-1}(u) = \alpha + (\beta - \alpha)u$.

g) Weibull(θ, τ): $f(x) = \dfrac{\tau (x/\theta)^\tau e^{-(x/\theta)^\tau}}{x}$ where $x > 0$, $\theta > 0$, and $\tau > 0$.
$F(x) = 1 - e^{-(x/\theta)^\tau}, \quad E(X^k) = \theta^k \Gamma(1 + k/\tau)$ for $k > -\tau$. The Weibull$(\theta, \tau = 1)$ RV is the Exponential$(\lambda = 1/\theta)$ RV. $F^{-1}(u) = \theta[-\ln(1-u)]^{1/\tau}$ for $0 < u < 1$.

The following are discrete distributions. Note: $p_k = P(X = k) = p(k)$.

h) Bernoulli(p) = binomial($n = 1, p$) the pmf $p(x) = p^x(1-p)^{1-x}$ for $x = 0, 1$. $E(X) = p$, $V(X) = p(1-p)$, $\phi(t) = [(1-p) + pe^t]$.

i) binomial(n, p): n is a (usually known) positive integer

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for} \quad \text{x} = 0, 1, \ldots, \text{n} \quad \text{where} \quad 0 < \text{p} < 1.$$

$E(X) = np$, $V(X) = np(1-p)$, $\phi(t) = [(1-p) + pe^t]^n$.

j) geometric(p): $p(x) = (1-p)^{x-1}p$ for $x = 1, 2, 3\ldots$ where $0 < p < 1$. $E(X) = 1/p$, $V(X) = \dfrac{1-p}{p^2}$, and $\phi(t) = \dfrac{pe^t}{1-(1-p)e^t}$.

k) Poisson(λ): $p(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ for $x = 0, 1, \ldots$, where $\lambda > 0$. $E(X) = \lambda = V(X)$, $\phi(t) = \exp(\lambda(e^t - 1))$.

Some properties of the gamma function follow.

i) $\Gamma(k) = (k-1)!$ for integer $k \geq 1$.

ii) $\Gamma(x+1) = x\ \Gamma(x)$ for $x > 0$.

iii) $\Gamma(x) = (x-1)\ \Gamma(x-1)$ for $x > 1$.

iv) $\Gamma(0.5) = \sqrt{\pi}$.

The **cumulative distribution function** (cdf) $F(x) = P(X \leq x)$. Then $F(-\infty) = 0, F(\infty) = 1$, and $F(x)$ is nondecreasing.

The probability density function (**pdf**) $f(x) = F'(x)$.

The **survival function** $S(x) = P(X > x)$. $S(-\infty) = 1, S(\infty) = 0$ and $S(x)$ is nonincreasing.

Let $F(x-) = P(X < x)$.

A **set** consists of distinct elements enclosed by *braces*, eg $\{1, 5, 7\}$.

The *empty set* $\emptyset$ is the set that contains no elements.

A is a subset of B, $A \subseteq B$, if every element in A is in B.

The **union** of A with $B = A \cup B$ is the set of all elements in A or B (or in both).

The **intersection** of A with $B = A \cap B = AB$ is the set of all elements in A and B.

If $A \cap B = \emptyset$, then A and B are (**mutually exclusive** or) **disjoint sets.**

The *complement* of A is $\overline{A} = A^c$, the set of elements in S but not in A.

The *sample space* S is the set of all possible outcomes of an experiment. A *sample point* E_i is a possible outcome. An *event* is a subset of S. A simple event is a set that contains exactly one element of S, eg $A = \{E_3\}$. A *discrete sample space* consists of a finite or countable number of outcomes.

The *relative frequency interpretation of probability* says that the probability of outcome (sample point) E_i is the proportion of times that E_i would occur if the experiment was repeated again and again infinitely often.

For **any event** A, $0 \leq P(A) \leq 1$.

Three axioms: $P(A) \geq 0, P(S) = 1$, and if $A_1, A_2, \ldots$ are pairwise mutually exclusive, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

1) **Common problem.** Use order to find S. Using a table to find S if two die are tossed or if a die is tossed twice and to find S if a coin is flipped 2, 3, or 4 times are typical examples.

The *sample point method* for finding the probability for event A says that if $S = \{E_1, ..., E_k\}$ then $0 \le P(E_i) \le 1$, $\sum_{i=1}^k P(E_i) = 1$, and $P(A) = \sum_{i:E_i \in A} P(E_i)$. That is, $P(A)$ is the sum of the probabilities of the sample points in A. If all of the outcomes E_i are *equally likely,* then $P(E_i) = 1/k$ and $P(A) =$ (number of outcomes in $A)/k$ if S contains k outcomes.

2) **Common Problem.** Leave the probabilities of some outcomes blank. See HW1 6.

3) **Common problem.** List all outcomes in S and use these outcomes to find $P(A)$.

The *multiplication principle* says that if there are n_1 ways to do a first task, n_2 ways to do a 2nd task, ..., and n_k ways to do a kth task, then the number of ways to perform the total act of performing the 1st task, then the second task, ..., then the kth task is $n_1 \cdot n_2 \cdot n_3 \cdots n_k$. Techniques for multiplication principle: a) use a slot for each task and write n_i above the ith task. There will be k slots, one for each task. b) use a tree diagram. Ex. There are n^r *samples of size r with replacement* taken from n objects.

A special case is the *number of permutations* (ordered arrangements using r of n distinct objects) $= P_r^n = n \cdot (n-1) \cdot (n-2) \cdots (n-r+1) = \dfrac{n!}{(n-r)!}$. The story problem has r slots and *order is important.* No object is allowed to be repeated in the arrangement. Typical questions include *how many ways to* "choose r people from n and arrange in a line," "to make r letter words with no letter repeated", "to make 7 digit phone numbers with no digit repeated." Key words include *order, no repeated* and *different.*

A special case of permutations is $P_n^n = n! = n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdots 4 \cdot 3 \cdot 2 \cdot 1 = n \cdot (n-1)! = n \cdot (n-1) \cdot (n-2)! = n \cdot (n-1) \cdot (n-2) \cdot (n-3)! = \cdots$. Typical problems include *sample of size r without replacement*, the number of ways to arrange n books, to arrange the letters in the word CLIPS (5!), etc.

4) **Common Problem:** Use the multiplication principle to find *how many* ways to perform the total task. The number of ways to answer k TF questions (2^k) and the number of ways to answer k multiple choice questions with n options (n^k) where $n = 4$ or 5, and the number of serial or license numbers are typical examples.

A *combination* is an unordered selection using r of n distinct objects. The *number of combinations* is the binomial coefficient $C_r^n = \binom{n}{r} = \dfrac{n!}{r!(n-r)!}$. This formula is used in story problems where *order is not important.* Key words include *committees, selecting* (eg 4 people from 10), *choose*, and *unordered.*

5) **Common problem.** Use the combination formula to solve a story problem.

6) **Common Problem.** Often the multiplication principle will be combined with combinations, permutations, powers and factorials.

7) **Common problem.** Use counting rules to find $P(A)$ when the outcomes in S are equally likely. Card problems are typical.

The *conditional probability* of A *given* B is $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ if $P(B) > 0$. Think of this probability as an experiment with sample space B instead of S. Key word: *given*.

8) **Common Problem.** You are given a table with i rows and j columns and asked to find conditional and unconditional probabilities. Find the row, column, and grand totals.

Two events A and B are **independent** if any of the following three conditions hold: i) $P(A \cap B) = P(A)P(B)$, ii) $P(A|B) = P(A)$, or iii) $P(B|A) = P(B)$. If *any of these conditions fails to hold,* then A and B are *dependent.*

9) **Common Problem.** Given some of $P(A), P(B), P(A \cap B)$, and $P(A|B)$, find $P(A|B)$ and whether A and B are independent.

10) **Common Problem.** Given a table as in 8), determine if a row event and a column event are independent.

Multiplication law. $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$. If A and B are ind., then $P(A \cap B) = P(A)P(B)$. If $A_1, A_2, ..., A_k$ are ind., then $P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1) \cdots P(A_k)$. In general $P(A_1 \cap A_2 \cap \cdots \cap A_k) =$
$P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)P(A_4|A_1 \cap A_2 \cap A_3) \cdots P(A_k|A_1 \cap A_2 \bigcap \cdots \cap A_{k-1})$.

Complement rule. $P(A) = 1 - P(\overline{A})$.

Additive law for disjoint events. If A and B are disjoint, then $P(A \cup B) = P(A) + P(B)$. If $A_1, ..., A_k$ are disjoint, then $P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1) + \cdots + P(A_k)$.

General Additive Rule. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

11) **Common Problem.** Given three of the 4 probabilities above, find the 4th. Variants: given $P(A)$ and $P(B)$ find $P(A \cup B)$ if A and B are disjoint or independent. 11) (continued) Use the addition rule to determine whether A and B are independent or disjoint.

Events $B_1, B_2, ..., B_k$ partition S if a) $B_i \cap B_j = \emptyset$ for $i \neq j$, b) $P(A_i) > 0$ for $i = 1, ..., k$, and c) $B_1 \cup B_2 \cup \cdots \cup B_k = S$. Often $k = 2$, and A and the complement $\overline{A}$ form a partition of S.

Let $B_1, B_2, ..., B_k$ partition S, and let A be an event in S, then
a) $P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_k)P(A|B_k) = \sum_{i=1}^k P(B_i)P(A|B_i)$ and
b) **Bayes' rule:** $P(B_j|A) = \dfrac{P(B_j \cap A)}{P(A)} = \dfrac{P(B_j)P(A|B_j)}{P(A)}$

$$= \frac{P(B_j)P(A|B_j)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_k)P(A|B_k)} = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^k P(B_i)P(A|B_i)}.$$

In particular, if $n = 2$, $P(E) = P(A)P(E|A) + P(\overline{A})P(E|\overline{A})$ and
$P(A|E) = \dfrac{P(A)P(E|A)}{P(A)P(E|A) + P(\overline{A})P(E|\overline{A})}$.

In a **Bayes' rule** story problem, 2 or more unconditional probabilities are given (or easy to find with the complement rule). Several conditional probabilities are also given (or easy to find with the complement rule). **Make a tree diagram** with the events corresponding to the unconditional events labelling the left branches and the events corresponding to the conditional probabilities labelling the right branches. Above the left branches place the unconditional probabilities and above the right branches place the conditional probabilities. You will be asked to find an unconditional right branch probability and to use Bayes' rule to find P(left branch | right branch).

Tips: the hard conditional probability, P(left branch | right branch), usually appears at the end of the story problem. This tells you how to label the left branches and the right branches of the tree. (The easy conditional probabilities, P(right branch | left branch), can also tell you how to label the tree.) The probabilities of the left branch sum to one. Each subtree of right branches has probabilities that sum to one. Occasionally you are asked to find both a P(right branch | left branch) (directly from the tree) and P(left branch | right branch) (using Bayes rule). **See HW2 7c** for how to do Bayes problems. Jar (or urn) problems are common.

12) **Common E1 problem:** Draw a tree diagram and use the tree to find the unconditional probability of a right branch event and use the tree and Bayes rule to find a conditional probability of a left branch event given a right branch event.

A *random variable* (RV) is a real valued function with a sample space as a domain. The *population* is the entire group of objects from which we want information. The *sample* is the part of the pop. actually examined.

The **distribution function** of any RV Y is $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

A RV is *discrete* if it can assume only a finite or countable number of distinct values. The collection of these probabilities is the *probability distribution* of the discrete RV. The *probability mass function* of a discrete RV Y is $p(y) = P(Y = y)$ where $0 \leq p(y) \leq 1$ and $\sum_{y:p(y)>0} p(y) = 1$.

13) **Common Problem.** The sample space of Y is $S_Y = \{y_1, y_2, ..., y_k\}$ and a table of y_k and $P(y_k)$ is given with one $P(y_k)$ omitted. Find the omitted $p(y_k)$ by using the fact that $\sum_{i=1}^k p(y_i) = p(y_1) + p(y_2) + \cdots + p(y_k) = 1$.

Let Y be a discrete RV with probability function $p(y)$. Then the *mean* or **expected value** of Y is $E(Y) = \mu = \displaystyle\sum_{y:p(y)>0} y\ p(y)$. If $g(Y)$ is a real valued function of Y, then $g(Y)$ is a random variable and $E[g(Y)] = \displaystyle\sum_{y:p(y)>0} g(y)\ p(y)$. The *variance* of Y is $V(Y) = E[(Y - E(Y))^2]$ and the *standard deviation* of Y is SD(Y) $= \sigma = \sqrt{V(Y)}$.

Short cut formula for variance. $V(Y) = E(Y^2) - (E(Y))^2$

If $S_Y = \{y_1, y_2, ..., y_k\}$ then $E(Y) = \displaystyle\sum_{i=1}^k y_i p(y_i) = y_1 p(y_1) + y_2 p(y_2) + \cdots + y_k p(y_k)$ and $E[g(Y)] = \displaystyle\sum_{i=1}^k g(y_i) p(y_i) = g(y_1)p(y_1) + g(y_2)p(y_2) + \cdots + g(y_k)p(y_k)$. Also $V(Y) = \displaystyle\sum_{i=1}^k (y_i - E(Y))^2 p(y_i) = (y_1 - E(Y))^2 p(y_1) + (y_2 - E(Y))^2 p(y_2) + \cdots + (y_k - E(Y))^2 p(y_k)$. Often using $V(Y) = E(Y^2) - (E(Y))^2$ is simpler where $E(Y^2) = y_1^2 p(y_1) + y_2^2 p(y_2) + \cdots + y_k^2 p(y_k)$.

14) **COMMON PROBLEM.** Given a table of y and $p(y)$, find $E[g(Y)]$ and the

standard deviation $\sigma = SD(Y)$.

$E(c) = c$, $E(cg(Y)) = cE(g(Y))$, and $E[\sum_{i=1}^k g_i(Y)] = \sum_{i=1}^k E[g_i(Y)]$ where c is any constant.

Suppose there are n independent identical trials and Y counts the number of successes and the $p =$ prob of success for any given trial. Then $Y \sim bin(n,p)$. Let D_i denote a S in the ith trial. Then

i) P(none of the n trials were successes) $= (1-p)^n = P(Y=0) = P(\overline{D}_1 \cap \overline{D}_2 \cap \cdots \cap \overline{D}_n)$.
ii) P(at least one of the trials was a success) $= 1-(1-p)^n = P(Y \geq 1) = 1 - P(Y=0)$ $= 1 - P(none) = P(\overline{\overline{D}_1 \cap \overline{D}_2 \cap \cdots \cap \overline{D}_n})$.
iii) P(all n trials were successes) $= p^n = P(Y=n) = P(D_1 \cap D_2 \cap \cdots \cap D_n)$.
iv) P(not all n trials were successes) $= 1-p^n = P(Y < n) = 1 - P(Y=n) = 1 -$ P(all).

Know: P(Y was at least k) $= P(Y \geq k)$ and P(Y at most k) $= P(Y \leq k)$.

15) **Common Problem.** Given a story problem, recognize that Y is bin(n,p), find E(Y), SD(Y), V(Y), P(Y = y), P(Y is at least j) $= p(j) + \cdots + p(n) = 1 - p(0) - \cdots - p(j-1)$, or P(Y is at most j) $= p(0) + \cdots + p(j) = 1 - p(j+1) - p(j+2) - \cdots - p(n)$.

16) **Common Problem.** Given that Y is a Poisson(λ) RV, find E(Y), V(Y), P(Y = y), P(Y is at least j), P(Y is at most j).

A RV Y is **continuous** if its distribution function $F(y)$ is continuous.

If Y is a continuous RV, then the **probability density function** (pdf) of Y is $f(y) = \frac{d}{dy}F(Y)$ wherever the derivative exists (in this class the derivative will exist everywhere except possibly for a finite number of points). If $f(y)$ is a pdf, then $f(y) \geq 0$ $\forall y$ and $\int_{-\infty}^{\infty} f(t)dt = 1$.

17) **Common Problem.** Find $f(y)$ from $F(y)$.

Fact: If Y has pdf $f(y)$, then $F(y) = \int_{-\infty}^{y} f(t)dt$.

18) **Common Problem.** Find $F(y)$ from $f(y)$.

19) **Common Problem.** Given that $f(y) = c\; g(y)$, find c.

Fact: If Y has pdf $f(y)$, then $P(a < Y < b) = P(a < Y \leq b) = P(a \leq Y < b) = P(a \leq Y \leq b) = \int_a^b f(y)dy = F(b) - F(a)$. Notation: $F(y-) = P(Y < y)$.

Fact: If Y has a probability function $p(y)$, then Y is discrete and $P(a < Y \leq b) = F(b) - F(a)$, but $P(a \leq Y \leq b) \neq F(b) - F(a)$.

20) **Common Problem.** Given the pdf $f(y)$, find $P(a < Y < b)$, etc.

If Y has pdf $f(y)$, then the *mean* of **expected value** of Y is $E(Y) = \int_{-\infty}^{\infty} y f(y)dy$ and $E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy$. $V(Y) = \int_{-\infty}^{\infty} (y - E[Y])^2 f(y)dy$.

Short cut formula: $V(Y) = E[Y^2] - (E[Y])^2$. (True for discrete and continuous RV's.) The standard deviation of Y is $SD(Y) = \sqrt{V(Y)}$.

21) **COMMON FINAL PROBLEM.** Given the probability function $p(y)$ if Y is discrete or given the pdf $f(y)$ if Y is continuous, find $E[Y]$, $V(Y)$, $SD(Y)$, and $E[g(Y)]$. The functions $g(y) = y$, $g(y) = y^2$, and $g(y) = e^{ty}$ are especially common.