Math 480 Exam 2 is Wed. Nov. 2. **You are allowed 11 sheets of notes and a calculator.** The exam emphasizes HW5-8, and Q5-8.

From the 1st exam:

The *conditional probability* of A *given* B is $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ if $P(B) > 0$.

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

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

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 $E(Y) = \sum_{i=1}^{k} y_i p(y_i) = y_1 p(y_1) + y_2 p(y_2) + \cdots + y_k p(y_k)$ and $E[g(y)] = \sum_{i=1}^{k} g(y_i) p(y_i) = g(y_1) p(y_1) + g(y_2) p(y_2) + \cdots + g(y_k) p(y_k)$. Also $V(Y) = \sum_{i=1}^{k} (y_i - E(Y))^2 p(y_i) = (y_1 - E(Y))^2 p(y_1) + (y_2 - E(Y))^2 p(y_2) + \cdots + (y_k - E(Y))^2 p(y_k)$.

Often using $V(Y) = E(Y^2) - (E(Y))^2$ is simpler where $E(Y^2) = y_1^2 p(y_1) + y_2^2 p(y_2) + \cdots + y_k^2 p(y_k)$.

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

If Y has pdf $f(y)$, then $\int_{-\infty}^{\infty} f(y)dy = 1$, $F(y) = \int_{-\infty}^{y} f(t)dt$ and $f(y) = F'(y)$ except at possibly countably many points, $E[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y) dy$, $P(a < Y < b) = F(b) - F(a) = \int_a^b f(y)dy$ where $<$ can be replaced by $\leq$.

$F(y) = P(Y \leq y)$. If Y has a pmf, $P(a < Y \leq b) = F(b) - F(a)$.

MATERIAL "NOT ON 1st EXAM"

22) Know how to use most of the RVs from the first page of the exam 1 review. (The Poisson, Binomial, and Weibull are less likely.)

23) The *support* of RV Y is the set $\{y : f(y) > 0\}$ or $\{y : p(y) > 0\}$. Formulas for $F(y)$, $f(y)$, and $p(y)$ are often given for the support or for the support plus the boundaries of the support (often for $(-\infty, b]$, $[a, b]$ or $[a, \infty)$).

Suppose that Y is a RV and that $E(Y) = \mu$ and standard deviation $\sqrt{V(Y)} = SD(Y) = \sigma$ exist. Then the **z-score** is $Z = \dfrac{Y - \mu}{\sigma}$. Note that $E(Z) = 0$, and $V(Z) = 1$.

24) Know how to do a **forwards calculation using the Z table** where $Z \sim N(0, 1)$. In the story problem you will be told that X is approximately normal with some mean and SD(X) or $V(X)$. You will be given one or two x^* values and asked **to find a probability** or proportion. Draw a line and mark down the mean and the x^* values. Standardize with $z^* = (x^* - \mu)/\sigma$, and sketch a Z curve (N(0,1) pdf). Show how the Z table is used. Then $P(X \leq x^*) = P(Z \leq z^*)$, $P(X > x^*) = 1 - P(Z \leq z^*)$ and $P(x_1^* < X < x_2^*) = P(Z \leq z_2^*) - P(Z \leq z_1^*)$. Note that $<$ can be replaced by $\leq$. Given a z^*, use the leftmost column and top row of the Z table. Intersect this row and column to get a 4 digit decimal $= P(Z \leq z^*)$. Note that $P(Z > 3.5) \approx 0 \approx P(Z < -3.5)$ and $P(Z < 3.5) \approx 1 \approx P(Z > -3.5)$. Also, $P(Z > z) = P(Z < -z)$. The normal pdf is symmetric about μ so the N(0,1) pdf is symmetric about 0 and bell shaped. See HW5

and Q5.

25) Know how to do a **backwards calculation using the Z table.** Here you are **given a probability and asked to find** one or two x^* values, often a percentile x_p where $P(X \leq x_p) = p$ if $X \sim N(\mu, \sigma^2)$. The Z table gives areas to the **left** of z^*. So if you are asked to find the top 5%, that is the same as finding the bottom 95%. If you are asked to find the bottom 25%, the Z table gives the correct value. If you are asked to find the two values containing the middle 95%, then 5% of the area is outside of the middle. Hence .025 area is to the left of $x^*(lo)$ and $.025 + .95 = .975$ area is to the left of $x^*(hi)$. The area to the left of x^*, is also the area to the left of z^*. Find the largest 4 digit number smaller than the desired area and the smallest 4 digit number larger than the desired area. These two numbers will be found in the middle of the Z table. Take the number closest to the desired area, and to find the corresponding z^*, examine the row and column containing the number. If there is a tie, average the two numbers to get z^*. Go along the row to the entry in the leftmost column of the Z table and go along the column to the top row of the Z table. For example, if your 4 digit number is .9750, $z^* = 1.96$. To get the corresponding x^*, use $x^* = \mu + \sigma z^*$. The 5th percentile has $z^* = -1.645$ and the 95th percentile has $z^* = 1.645$ because there is a tie. See HW5, Q5.

Let Y_1 and Y_2 be discrete random variables. Then the **joint probability mass function** $p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2)$ and is often given by a table.

The function $p(y_1, y_2)$ is a joint pmf if $p(y_1, y_2) \geq 0$, $\forall y_1, y_2$ and if
$$\sum_{(y_1,y_2):p(y_1,y_2)>0} p(y_1, y_2) = 1.$$

The **joint cdf** of any two random variables Y_1 and Y_2 is
$F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2), \forall y_1, y_2.$

Let Y_1 and Y_2 be continuous random variables. Then the **joint probability density function** $f(y_1, y_2)$ satisfies $F(y_1, y_2) = \int_{-\infty}^{y_2} \int_{-\infty}^{y_1} f(t_1, t_2) dt_1 dt_2 \quad \forall y_1, y_2$.

The function $f(y_1, y_2)$ is a joint pdf if $f(y_1, y_2) \geq 0, \forall y_1, y_2$ and
$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1.$

$P(a_1 < Y_1 < b_1, a_2 < Y_2 < b_2) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(y_1, y_2) dy_1 dy_2$

$F(y_1, ..., y_n) = P(Y_1 \leq y_1, ..., Y_n \leq y_n)$. In the discrete case, the multivariate probability function is $p(y_1, ..., y_n) = P(Y_1 = y_1, ..., Y_n = y_n)$. In the continuous case, $f(y_1, ..., y_n)$ is a joint pdf if $F(y_1, ..., y_n) = \int_{-\infty}^{y_n} ... \int_{-\infty}^{y_1} f(t_1, ..., t_n) dt_1 \cdots dt_n$.

26) **Common Problem.** If $p(y_1, y_2)$ is given by a table, the marginal probability functions are found from the row sums and column sums and the conditional probability functions are found with the above formulas.

27) **COMMON FINAL PROBLEM.** Given the joint pdf $f(y_1, y_2) = kg(y_1, y_2)$ on its support, find k, find the marginal pdf's $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ and find the conditional pdf's $f_{Y_1|Y_2=y_2}(y_1|y_2)$ and $f_{Y_2|Y_1=y_1}(y_2|y_1)$.

Often using **symmetry** helps.

The *support* of the conditional pdf can depend on the 2nd variable. For example, the support of $f_{Y_1|Y_2=y_2}(y_1|y_2)$ could have the form $0 \leq y_1 \leq y_2$.

Double Integrals. If the region of integration Ω is bounded on top by the function $y_2 = \phi_T(y_1)$, on the bottom by the function $y_2 = \phi_B(y_1)$ and to the left and right by the lines $y_1 = a$ and $y_2 = b$ then $\int\int_\Omega f(y_1, y_2) dy_1 dy_2 = \int\int_\Omega f(y_1, y_2) dy_2 dy_2 =$

$$\int_a^b [\int_{\phi_B(y_1)}^{\phi_T(y_1)} f(y_1, y_2) dy_2] dy_1.$$

Within the inner integral, treat y_2 as the variable, anything else, including y_1, is treated as a constant.

If the region of integration Ω is bounded on the left by the function $y_1 = \psi_L(y_2)$, on the right by the function $y_1 = \psi_R(y_2)$ and to the top and bottom by the lines $y_2 = c$ and $y_2 = d$ then $\int\int_\Omega f(y_1, y_2) dy_1 dy_2 = \int\int_\Omega f(y_1, y_2) dy_2 dy_2 =$

$$\int_c^d [\int_{\psi_L(y_2)}^{\psi_R(y_2)} f(y_1, y_2) dy_1] dy_2.$$

Within the inner integral, treat y_1 as the variable, anything else, including y_2, is treated as a constant.

The **support** of continuous random variables Y_1 and Y_2 is where $f(y_1, y_2) > 0$. The support (plus some points on the boundary of the support) is generally given by one to three inequalities such as $0 \leq y_1 \leq 1$, $0 \leq y_2 \leq 1$, and $0 \leq y_1 \leq y_2 \leq 1$. For each variable, set the inequalities to equalities to get boundary lines. For example $0 \leq y_1 \leq y_2 \leq 1$ yields 5 lines: $y_1 = 0, y_1 = 1$, $y_2 = 0, y_2 = 1$, and $y_2 = y_1$. Generally y_2 is on the vertical axis and y_1 is on the horizontal axis for pdf's.

To determine the **limits of integration**, examine the **dummy variable used in the inner integral**, say dy_1. Then within the region of integration, draw a line parallel to the same (y_1) axis as the dummy variable. The limits of integration will be functions of the other variable (y_2), never of the dummy variable (dy_1).

If Y_1 and Y_2 are discrete RV's with joint probability function $p(y_1, y_2)$, then the **marginal pmf for** Y_1 is

$$p_{Y_1}(y_1) = \sum_{y_2} p(y_1, y_2)$$

where y_1 is held fixed. The **marginal pmf for** Y_2 is

$$p_{Y_2}(y_2) = \sum_{y_1} p(y_1, y_2)$$

where y_2 is held fixed. The **conditional pmf of** Y_1 **given** $Y_2 = y_2$ is

$$p_{Y_1|Y_2=y_2}(y_1|y_2) = \frac{p(y_1, y_2)}{p_{Y_2}(y_2)}.$$

The **conditional pmf of** Y_2 **given** $Y_1 = y_1$ is

$$p_{Y_2|Y_1=y_1}(y_2|y_1) = \frac{p(y_1, y_2)}{p_{Y_1}(y_1)}.$$

If Y_1 and Y_2 are continuous RV's with joint pdf $f(y_1, y_2)$, then the **marginal probability density function for** Y_1 is

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_2 = \int_{\phi_B(y_1)}^{\phi_T(y_1)} f(y_1, y_2) dy_2$$

where y_1 is held fixed (get the region of integration, draw a line parallel to the y_2 axis and use the functions $y_2 = \phi_B(y_1)$ and $y_2 = \phi_T(y_1)$ as the lower and upper limits of integration). The **marginal probability density function for** Y_2 is

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 = \int_{\psi_L(y_2)}^{\psi_R(y_2)} f(y_1, y_2) dy_1$$

where y_2 is held fixed (get the region of integration, draw a line parallel to the y_1 axis and use the functions $y_1 = \psi_L(y_2)$ and $y_1 = \psi_R(y_2)$ as the lower and upper limits of integration). The **conditional probability density function of** Y_1 **given** $Y_2 = y_2$ is

$$f_{Y_1|Y_2=y_2}(y_1|y_2) = \frac{f(y_1, y_2)}{f_{Y_2}(y_2)}$$

provided $f_{Y_2}(y_2) > 0$. The **conditional probability density function of** Y_2 **given** $Y_1 = y_1$ is

$$f_{Y_2|Y_1=y_1}(y_2|y_1) = \frac{f(y_1, y_2)}{f_{Y_1}(y_1)}$$

provided $f_{Y_1}(y_1) > 0$.

Random variables Y_1 and Y_2 are **independent** if any one of the following conditions holds.
i) $F(y_1, y_2) = F_{Y_1}(y_1) F_{Y_2}(y_2) \quad \forall y_1, y_2$.
ii) $p(y_1, y_2) = p_{Y_1}(y_1) p_{Y_2}(y_2) \quad \forall y_1, y_2$.
iii) $f(y_1, y_2) = f_{Y_1}(y_1) f_{Y_2}(y_2) \quad \forall y_1, y_2$.
Otherwise, Y_1 and Y_2 are *dependent*.

If $Y_1, Y_2, ..., Y_n$ are independent if $\forall y_1, y_2, ..., y_n$:
i) $F(y_1, y_2, ..., y_n) = F_{Y_1}(y_1) F_{Y_2}(y_2) \cdots F_{Y_n}(y_n)$
ii) $p(y_1, y_2, ..., y_n) = p_{Y_1}(y_1) p_{Y_2}(y_2) \cdots p_{Y_n}(y_n)$ or
iii) $f(y_1, y_2, ..., y_n) = f_{Y_1}(y_1) f_{Y_2}(y_2) \cdots f_{Y_n}(y_n)$. Otherwise, the Y_i are dependent.

Two RV's Y_1 and Y_2 are dependent if their support is not a cross product of the support of Y_1 with the support of Y_2. (A rectangular support is an important special case.) If the support is a cross product, another test must be used to determine whether Y_1 and Y_2 are independent or dependent.

For continuous Y_1 and Y_2, then Y_1 and Y_2 are independent iff $f(y_1, y_2) = g(y_1) h(y_2)$ on **cross product support** where g is a positive function of y_1 alone and h is a positive function of y_2 alone. Or use $f(y_1, y_2) = g(y_1) h(y_2)$ for nonnegative h and g for all y_1 and y_2 (not just the cross product support).

To check whether discrete Y_1 and Y_2 (with rectangular support) are independent given a 2 by 2 table, find the row and column sums and check whether $p(y_1, y_2) \neq p_{Y_1}(y_1) p_{Y_2}(y_2)$

for **some entry** (y_1, y_2). Then Y_1 and Y_2 are dependent. If $p(y_1, y_2) = p_{Y_1}(y_1)p_{Y_2}(y_2)$ for *all table entries, then* Y_1 *and* Y_2 *are independent.*

28) **Common Problem.** Determine whether Y_1 and Y_2 are independent or dependent.

Suppose that (Y_1, Y_2) are jointly continuous with joint pdf $f(y_1, y_2)$.
Then the **expectation** $E[g(Y_1, Y_2)] = \int_{\chi_1} \int_{\chi_2} g(y_1, y_2) f(y_1, y_2) dy_2 dy_1 =$
$\int_{\chi_2} \int_{\chi_1} g(y_1, y_2) f(y_1, y_2) dy_1 dy_2$ where χ_i are the limits of integration for dy_i.

In particular, $E(Y_1 Y_2) = \int_{\chi_1} \int_{\chi_2} y_1 y_2 f(y_1, y_2) dy_2 dy_1 = \int_{\chi_2} \int_{\chi_1} y_1 y_2 f(y_1, y_2) dy_1 dy_2$

If g is a function of Y_i but not of Y_j, find the marginal for Y_i : If $g(Y_1)$ is a function of Y_1 but not of Y_2, then $E[g(Y_1)] = \int_{\chi_1} \int_{\chi_2} g(y_1) f(y_1, y_2) dy_2 dy_1 = \int_{\chi_1} g(y_1) f_{Y_1}(y_1) dy_1$. (**Usually finding the marginal is easier than doing the double integral**.) Similarly, $E[g(Y_2)] = \int_{\chi_2} g(y_2) f_{Y_2}(y_2) dy_2$.

In particular, $E(Y_1) = \int_{\chi_1} y_1 f_{Y_1}(y_1) dy_1$, and $E(Y_2) = \int_{\chi_2} y_2 f_{Y_2}(y_2) dy_2$.

Suppose that (Y_1, Y_2) are jointly discrete with joint probability function $p(y_1, y_2)$. Then the **expectation** $E[g(Y_1, Y_2)] = \sum_{y_2} \sum_{y_1} g(y_1, y_2) p(y_1, y_2) = \sum_{y_1} \sum_{y_2} g(y_1, y_2) p(y_1, y_2)$.

In particular, $E[Y_1 Y_2] = \sum_{y_2} \sum_{y_1} y_1 y_2 p(y_1, y_2)$.

If g is a function of Y_i but not of Y_j, find the marginal for Y_i. If $g(Y_1)$ is a function of Y_1 but not of Y_2, then $E[g(Y_1)] = \sum_{y_2} \sum_{y_1} g(y_1) p(y_1, y_2) = \sum_{y_1} g(y_1) p_{Y_1}(y_1)$. (**Usually finding the marginal is easier than doing the double summation**.) Similarly, $E[g(Y_2)] = \sum_{y_2} g(y_2) p_{Y_2}(y_2)$.

In particular, $E(Y_1) = \sum_{y_1} y_1 p_{Y_1}(y_1)$ and $E(Y_2) = \sum_{y_2} y_2 p_{Y_2}(y_2)$.

The **covariance** of Y_1 and Y_2 is $Cov(Y_1, Y_2) = E[(Y_1 - E(Y_1))(Y_2 - E(Y_2))]$.

Short cut formula: $Cov(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1) E(Y_2)$.

Let Y_1 and Y_2 **be independent random variables**. If g is a function of Y_1 alone and h is a function of Y_2 alone, then $g(Y_1)$ and $h(Y_2)$ are independent random variables and $E[g(Y_1) h(Y_2)] = E[g(Y_1)] E[h(Y_2)]$ if the expectations exist.
In particular, $E[Y_1 Y_2] = E[Y_1] E[Y_2]$.

Know: Let Y_1 and Y_2 be independent random variables. Then $Cov(Y_1, Y_2) = 0$.

The converse is false: it is possible that $Cov(Y_1, Y_2) = 0$ but Y_1 and Y_2 are dependent.

29) **COMMON FINAL PROBLEM.** If $p(y_1, y_2)$ is given by a table, determine whether Y_1 and Y_2 are independent or dependent, find the marginal probability functions $p_{Y_1}(y_1)$ and $p_{Y_2}(y_2)$ and find the conditional probability function's $p_{Y_1|Y_2=y_2}(y_1|y_2)$ and $p_{Y_2|Y_1=y_1}(y_2|y_1)$. Also find $E(Y_1), E(Y_2), V(Y_1), V(Y_2), E(Y_1 Y_2)$ and $Cov(Y_1, Y_2)$.

30) **COMMON FINAL PROBLEM.** Given the joint pdf $f(y_1, y_2) = k g(y_1, y_2)$ on its support, find k, find the marginal pdf's $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ and find the conditional pdf's $f_{Y_1|Y_2=y_2}(y_1|y_2)$ and $f_{Y_2|Y_1=y_1}(y_2|y_1)$. Also determine whether Y_1 and Y_2 are independent or dependent, and find $E(Y_1), E(Y_2), V(Y_1), V(Y_2), E(Y_1 Y_2)$ and $Cov(Y_1, Y_2)$. If $Cov(Y_1, Y_2) \neq 0$, or if the support is not a cross product, then Y_1 and Y_2 are dependent. If $Cov(Y_1, Y_2) = 0$ and if the support is a cross product, you cannot tell whether Y_1 and Y_2 are dependent or not. In this case if you can show that $f(y_1, y_2) = g(y_1) h(y_2)$ on its cross product support or that $f(y_1, y_2) = f_{Y_1}(y_1) f_{Y_2}(y_2)$, then Y_1 and Y_2 are independent,

otherwise Y_1 and Y_2 are dependent.

Often using **symmetry** helps.

$E(c) = c$, $E[g_1(Y_1,Y_2) + \cdots + g_k(Y_1,Y_2)] = E[g_1(Y_1,Y_2)] + \cdots + E[g_k(Y_1,Y_2)]$. In particular, $E[aY_1 + bY_2] = aE[Y_1] + bE[Y_2]$.

Know: Let a be any constant and let Y be a RV. Then $E[aY] = aE[Y]$ and $V(aY) = a^2 V(Y)$.

Let $Y_1, ..., Y_n$, and $X_1, ..., X_m$ be random variables. Let $U_1 = \sum_{i=1}^n a_i Y_i$ and $U_2 = \sum_{i=1}^m b_i X_i$ for constants $a_1, ..., a_n, b_1, ..., b_n$. Then $E(U_1) = \sum_{i=1}^n a_i E(Y_i)$, $V(U_1) = \sum_{i=1}^n a_i^2 V(Y_i) + 2 \sum\sum_{i<j} a_i a_j Cov(Y_i, Y_j)$ (so i goes from 1 to $n-1$ and j from $i+1$ to n) and $Cov(U_1, U_2) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(Y_i, X_j)$.

31) **Common problem (Not in Text):** Find the pmf of $Y = t(X)$ and the sample space $\mathcal{Y}$ given the pmf $p_X(x)$ of X in a table. Step 1) Find $y = t(x)$ for each value of x. Step 2) Collect $x : t(x) = y$, and sum the corresponding probabilities:
$p_Y(y) = \sum_{x:t(x)=y} p_X(x)$, and table the result.
For example, if $Y = X^2$ and $p_X(-1) = 1/3, p_X(0) = 1/3$, and $p_X(1) = 1/3$, then $p_Y(0) = 1/3$ and $p_Y(1) = 2/3$.

32) **Common problem (Not in Text), the method of transformations:** Find the pdf of $Y = t(X)$ and the sample space $\mathcal{Y}$ given the pdf of X where t is increasing or decreasing:

$$f_Y(y) = f_X(t^{-1}(y)) \left| \frac{dt^{-1}(y)}{dy} \right|$$

for $y \in \mathcal{Y}$. To be useful, this formula should be simplified as much as possible. To find the support $\mathcal{Y}$ of $Y = t(X)$ if the support of X is $\mathcal{X} = [a, b]$, plug in $t(x)$ and find the minimum and maximum value on $[a, b]$. A graph can help. If t is an increasing function, then $\mathcal{Y} = [t(a), t(b)]$. If t is an decreasing function, then $\mathcal{Y} = [t(b), t(a)]$.

Tips: a) The pdf of Y will often be that of a gamma RV. In particular, the pdf of Y is often the pdf of an exponential(λ) RV.
b) To find the inverse function $x = t^{-1}(y)$, solve the equation $y = t(x)$ for x.
c) The log transformation is often used. Know how to sketch $\log(y)$ and e^y for $y > 0$. Recall that in this class, $\log(y)$ is the natural logarithm of y.

The method of distribution functions: Suppose that the distribution function $F_X(x)$ is known, $Y = t(X)$, and both X and Y have pdfs.
a) If t is an increasing function then, $F_Y(y) = P(Y \le y) = P(t(X) \le y) =$
$P(X \le t^{-1}(y)) = F_X(t^{-1}(y))$ for $y \in \mathcal{Y}$.
b) If t is a decreasing function then, $F_Y(y) = P(Y \le y) = P(t(X) \le y) =$
$P(X \ge t^{-1}(y)) = 1 - P(X < t^{-1}(y)) = 1 - F_X(t^{-1}(y))$ for $y \in \mathcal{Y}$.
c) The special case $Y = X^2$ is important. If the support of X is positive, use a). If the support of X is negative, use b). If the support of X is $(-a, a)$ (where $a = \infty$ is allowed), then $F_Y(y) = P(Y \le y) =$

$P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) =$

$$\int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \text{ for } 0 \le \text{y} \le \text{a}^2$$

and

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})]$$

for $0 \le y \le a^2$.

33) **Common Problem:** Given two independent RV's X and Y, written $X \perp\!\!\!\perp Y$ find $E(aX \pm bY) = aE(X) \pm bE(Y)$ and $V(aX \pm bY) = a^2V(X) + b^2V(Y)$.

34) The **moment generating function** (mgf) of a random variable Y is $m(t) = \phi(t) = E[e^{tY}]$. If Y is discrete, then $\phi(t) = \sum_y e^{ty}p(y)$, and if Y is continuous, then $\phi(t) = \int_{-\infty}^{\infty} e^{ty}f(y)dy$. The **kth moment** of Y is $E[Y^k]$. Given the mgf $\phi(t)$ exists for $|t| < b$ for some constant $b > 0$, find the kth derivative $\phi^{(k)}(t)$. Then $E[Y^k] = \phi^{(k)}(0)$. In particular, $E[Y] = \phi'(0)$ and $E[Y^2] = \phi''(0)$.

Derivatives. The **product rule** is $(f(y)g(y))' = f'(y)g(y) + f(y)g'(y)$. The **quotient rule** is $\left(\frac{n(y)}{d(y)}\right)' = \frac{d(y)n'(y) - n(y)d'(y)}{[d(y)]^2}$. Know how to find 2nd, 3rd, etc derivatives. The **chain rule** is $[f(g(y))]' = [f'(g(y))][g'(y)]$. Know the derivative of ln y = $\log(y)$ and e^y and know the chain rule with these functions. Know the derivative of y^k.

35) The **probability generating function** (pgf) of a random variable X is $P_X(z) = E[z^X]$. If X is discrete, then $P_X(z) = \sum_x z^x p(x)$, and if X is continuous, then $P_X(z) = \int_{-\infty}^{\infty} z^x f(x)dx$. If the pgf $P_X(z)$ exists for $z \in (1-\epsilon, 1+\epsilon)$ for some constant $\epsilon > 0$, find the kth derivative $P_X^{(k)}(z)$. Then $E[X(X-1)\cdots(X-k+1)] = P_X^{(k)}(1)$ where the product has k terms. In particular, $E[X] = P_X'(1)$ and $E[X^2 - X] = E(X^2) - E(X) = P_X''(1)$.

36) $\phi_X(t) = P_X(e^t)$ and $P_X(z) = \phi_X(\log(z))$.

37) Let $S_n = \sum_{i=1}^n X_i$ where the X_i are independent with mgf $\phi_{X_i}(t)$ and pgf $P_{X_i}(z)$. The mgf of S_n is $\phi_{S_n}(t) = \prod_{i=1}^n \phi_{X_i}(t) = \phi_{X_1}(t)\phi_{X_2}(t)\cdots\phi_{X_n}(t)$. The pgf of S_n is $P_{S_n}(z) = \prod_{i=1}^n P_{X_i}(z) = P_{X_1}(z)P_{X_2}(z)\cdots P_{X_n}(z)$.

Tips: a) in the product, anything that does not depend on the product index i is treated as a constant. b) $\exp(a) = e^a$ and $\log(y) = \ln(y) = \log_e(y)$ is the **natural logarithm**. c) $\prod_{i=1}^n a^{b\theta_i} = a^{\sum_{i=1}^n b\theta_i}$. In particular, $\prod_{i=1}^n \exp(b\theta_i) = \exp(\sum_{i=1}^n b\theta_i)$. d) $\sum_{i=1}^n b = nb$. e) $\prod_{i=1}^n a = a^n$.

X has a negative binomial distribution, $X \sim NB(k,p)$ if the pmf of X is

$$p(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k} \text{ for } \text{x} = \text{k}, \text{k}+1, \text{k}+2, \ldots \text{ where } 0 < \text{p} < 1$$

and k is a positive integer. Take $p(k) = p^k$. $E(X) = k/p$, $V(X) = k(1-p)/p^2$, $\phi(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^k$. If $X \sim NB(k=1,p)$, then $X \sim geom(p)$.

38) Assume the X_i are independent.
a) If $X_i \sim N(\mu_i, \sigma_i^2)$, with support $(-\infty, \infty)$, then $\sum_{i=1}^n X_i \sim N(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$, and $\sum_{i=1}^n (a_i + b_i X_i) \sim N(\sum_{i=1}^n (a_i + b_i \mu_i), \sum_{i=1}^n b_i^2 \sigma_i^2)$. Here a_i and b_i are fixed constants. Thus if $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, then $\overline{X} \sim N(\mu, \sigma^2/n)$.
b) If $X_i \sim G(\alpha_i, \lambda)$, then $\sum_{i=1}^n X_i \sim G(\sum_{i=1}^n \alpha_i, \lambda)$. Note that the X_i have the same λ, and if $\alpha_i \equiv \alpha$, then $\sum_{i=1}^n \alpha = n\alpha$. G stands for Gamma.
c) If $X_i \sim EXP(\lambda) \sim G(1, \lambda)$, then $\sum_{i=1}^n X_i \sim G(n, \theta)$.
d) If $X_i \sim \chi^2_{k_i} \sim G\left(\frac{k_i}{2}, 1/2\right)$, then $\sum_{i=1}^n X_i \sim \chi^2_{\sum_{i=1}^n k_i}$. If $k_i \equiv k$, then $\sum_{i=1}^n k = nk$.
e) If $X_i \sim$ Poisson(λ_i) then $\sum_{i=1}^n X_i \sim$ Poisson$(\sum_{i=1}^n \lambda_i)$. Note that if $\lambda_i \equiv \lambda$, then $\sum_{i=1}^n \lambda = n\lambda$.
f) If $X_i \sim bin(k_i, p)$, then $\sum_{i=1}^n X_i \sim bin(\sum_{i=1}^n k_i, p)$. Note that the X_i have the same p, and if $k_i \equiv k$, then $\sum_{i=1}^n k = nk$.
g) Let NB stand for negative binomial. If $X_i \sim NB(k_i, p)$, then $\sum_{i=1}^n X_i \sim NB(\sum_{i=1}^n k_i, p)$. Note that the X_i have the same p, and if $k_i \equiv k$, then $\sum_{i=1}^n k = nk$.
h) Let $X_i \sim geom(\beta) \sim NB(1, p)$. Then $\sum_{i=1}^n X_i \sim NB(n, p)$.

39) i) Given $\phi_X(t)$ or $P_X(t)$, use 34) and 35) to find $E(X)$, $E(X^2)$, or $E(X^2) - E(X)$. Then find $V(X)$ or $SD(X) = \sqrt{V(X)}$.

ii) Given a table for the pmf $p_X(x)$, find the mgf $\phi(t) = \phi_X(t) = \sum_x e^{tx} p_X(x)$, or the pgf $P_X(z) = \sum_x z^x p_X(x)$.

iii) Given ϕ_X or P_X as in ii), find the pmf $p_X(x)$.

iv) Given a brand name ϕ_X find the parameters of the brand name RV X.

40) Markov's inequality: If $E(X)$ exists and $X \geq 0$ in that the support of $X \subseteq [0, \infty)$, then for any constant $a > 0$, $P(X \geq a) \leq \frac{E(X)}{a}$.

41) Chebyshev's inequality: If $E(X) = \mu$ and $V(X) = \sigma^2$, then for any constant $k > 0$, $P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$. Also, $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$ so $P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$.

42) Strong Law of Large Numbers (SLLN): Let $X_1, X_2, ...$ be iid with $E(X_i) = \mu$. Then $\overline{X} \to \mu$ as $n \to \infty$.

43) Central Limit Theorem (**CLT**): Let $Y_1, ..., Y_n$ be iid with $E(Y) = \mu$ and $V(Y) = \sigma^2$. Let the sample mean $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$. Then

$$\sqrt{n}(\overline{Y}_n - \mu) \xrightarrow{D} N(0, \sigma^2).$$

Hence

$$\sqrt{n}\left(\frac{\overline{Y}_n - \mu}{\sigma}\right) = \sqrt{n}\left(\frac{\sum_{i=1}^n Y_i - n\mu}{n\sigma}\right) = \left(\frac{\overline{Y}_n - \mu}{\sigma/\sqrt{n}}\right) = \left(\frac{\sum_{i=1}^n Y_i - n\mu}{\sqrt{n}\sigma}\right) \xrightarrow{D} N(0, 1).$$

The notation $X \sim Y$ means that the random variables X and Y have the same distribution. The notation $Y_n \xrightarrow{D} X$ means that for large n we can approximate the cdf of Y_n by the cdf of X. The distribution of X is the limiting distribution or asymptotic distribution of Y_n, and the limiting distribution does not depend on n. For the CLT, notice that

$$Z_n = \sqrt{n}\left(\frac{\overline{Y}_n - \mu}{\sigma}\right) = \left(\frac{\overline{Y}_n - \mu}{\sigma/\sqrt{n}}\right)$$

is the z–score of $\overline{Y}$ and

$$Z_n = \left(\frac{\sum_{i=1}^n Y_i - n\mu}{\sqrt{n}\sigma} \right)$$

is the z–score of $\sum_{i=1}^n Y_i$. If $Z_n \xrightarrow{D} N(0,1)$, then the notation $Z_n \approx N(0,1)$, also written as $Z_n \sim AN(0,1)$, means approximate the cdf of Z_n by the standard normal cdf. Similarly, the notation

$$\overline{Y}_n \approx N(\mu, \sigma^2/n),$$

also written as $\overline{Y}_n \sim AN(\mu, \sigma^2/n)$, means approximate the cdf of $\overline{Y}_n$ as if $\overline{Y}_n \sim N(\mu, \sigma^2/n)$. Note that $U = U_n = \sum_{i=1}^n Y_i \approx N(n\mu, n\sigma^2)$ if the Y_i are iid. Note that the approximate distribution, unlike the limiting distribution, does depend on n. Use the limiting distribution or approximate distribution to find probabilities and percentiles.

44) **Common Problem.** Perform a **forwards calculation** for $\overline{Y}$ using the normal table. In the story problem you will be told that $Y_1, ..., Y_n$ are iid with some mean μ and standard deviation σ (or variance σ^2). You will be told that "the CLT holds" or that the Y_i are "approximately normal". You will be asked to find the probability that the sample mean is greater than a or less than b or between a and b. That is, find $P(\overline{Y} > a)$ $P(\overline{Y} < b)$ or $P(a < \overline{Y} < b)$ (the strict inequalities $(<, >)$ may be replaced with nonstrict inequalities $(\leq, \geq)$). Call a and b "ybar values."
Step 0) Find $\mu_{\overline{Y}} = \mu$ and $\sigma_{\overline{Y}} = \sigma/\sqrt{n}$.
Step 1) Draw the $\overline{Y}$ picture with μ and the "ybar values" labeled.
Step 2) Find the z-score for each "ybar value", eg $z = \dfrac{a - \mu}{\sigma/\sqrt{n}}$.
Step 3) Draw a z-picture (sketch a N(0,1) curve and shade the appropriate area).
Step 4) Use the standard normal table to find the appropriate probability.

The CLT is what allows one to perform forwards calculations with $\overline{Y}$. How large should n be to use the CLT? i) $n \geq 1$ for Y_i iid normal. ii) $n \geq 5$ for Y_i iid approximately normal. iii) If the Y_i are iid from a **highly skewed distribution**, do not use the normal approximation (forwards calculation) if $n \leq 29$. iv) If $n > 100$, usually the CLT will hold in this class.

45) **Common Problem (Not in Text).** You are told that the Y_i are iid from a highly skewed distribution and that the sample size $n \leq 29$. You are asked to perform a forwards calculation such as $P(\overline{Y} > a)$ if possible. **Solution:** not possible n is too small for the CLT to apply.

46) **Common Problem.** Perform a **forwards calculation** for $\sum_{i=1}^n Y_i$ using the normal table if the Y_i are iid. Step 0) Find $\mu_{\sum Y_i} = n\mu$ and $\sigma_{\sum Y_i} = \sqrt{n}\sigma$.
Step 1) Draw the $\sum_{i=1}^n Y_i$ picture with $n\mu$ and the "sum values" labeled.
Step 2) Find the z-score for each "sum value", eg $z = \dfrac{a - n\mu}{\sqrt{n}\sigma}$.
Step 3) Draw a z-picture (sketch a N(0,1) curve and shade the appropriate area).
Step 4) Use the standard normal table to find the appropriate probability.

47) Think of $W \sim X|Y = y$. Then $X|Y$ is a family of random variables. If $E(X|Y = y) = m(y)$, then the random variable $E(X|Y) = m(Y)$. Similarly if $V(X|Y = y) = v(y)$, then the random variable $V(X|Y) = v(Y) = E(X^2|Y) - [E(X|Y)]^2$.

48) Assume all relevant expectations exist. Then iterated expectations or the conditional mean formula is $E(X) = E[E(X|Y)] = E_Y[E_{X|Y}(X|Y)]$. The conditional variance formula is $V(X) = E[V(X|Y)] + V[E(X|Y)]$.

49) Let N be a counting RV with support $\subseteq \{0, 1, 2, ...\}$. Let $N \perp\!\!\!\perp X_i$ where the X_i are independent, $E(X_i) = E(X)$ and $V(X_i) = V(X)$. Let $S_N = X_1 + X_2 + \cdots + X_N = \sum_{i=1}^N X_i$. Then $E(S_N) = E(N)E(X)$ and $V(S_N) = V(X)E(N) + [E(X)]^2 V(N)$. If $N = 0$, then $S_N = 0$. $S = S_N$ is a compound RV and the distribution of N is the compounding distribution.

End probability, start stochastic processes.

50) A stochastic process $\{X(t) : t \in \tau\}$ is a collection of random variables where the set τ is often $[0, \infty)$. Often t is time and the random variable $X(t)$ is the state of the process at time t.

51) A stochastic process $\{X(t) : t \in \{1, 2, ...\}\}$ is a *white noise* if $X_1, ..., X_t, ...$ are iid with $E(X_i) = 0$ and $V(X_i) = \sigma^2$.

52) A stochastic process $\{Y(t) : t \in \{1, 2, ...\}\}$ is a *random walk* if $Y(t) = Y_t = Y_{t-1} + e_t$ where the e_t are iid and $Y_0 = y_0$ is a constant. Then $Y_t = Y_{t-2} + e_{t-1} + e_t = Y_{t-j} + e_{t-j+1} + \cdots + e_t = y_0 + e_1 + e_2 + \cdots + e_t = y_0 + \sum_{i=1}^t e_i$ where $\sum_{i=1}^t e_i$ is known as a cumulative sum. If $E(e) = \delta$ and $V(e) = \sigma^2$, then $E(Y_t) = y_o + t\delta$ and $V(Y_t) = t\sigma^2$.

Poisson Processes

53) A stochastic process $\{N(t) : t \geq 0\}$ is a counting process if $N(t)$ counts the total number of events that occurred in time interval $(0, t]$. If $0 < t_1 < t_2$, then the random variable $N(t_2) - N(t_1)$ counts the number of events that occurred in interval $(t_1, t_2]$.

54) $N(t)$ is said to possess independent increments if the number of events that occur in disjoint time intervals are independent. Hence if $0 < t_1 < t_2 < t_3 < \cdots < t_k$, then the RVs $N(t_1) - N(0), N(t_2) - N(t_1), ..., N(t_k) - N(t_{k-1})$ are independent.

55) $N(t)$ is said to possess stationary increments if the distribution of events that occur in any time interval depends only on the length of the time interval.

56) A counting process $\{N(t) : t \geq 0\}$ is a *Poisson process with rate* λ for $\lambda > 0$ if i) $N(0) = 0$, ii) the process has independent increments, iii) the number of events in any interval of length t has a Poisson (λt) distribution with mean λt.

57) Hence the Poisson process $N(t)$ has stationary increments, the number of events in $(s, s+t]$ = the number of events in $(s, s+t)$, and for all $t, s \geq 0$, the RV $D(t) = N(t+s) - N(s) \sim$ Poisson (λt). In particular, $N(t) \sim$ Poisson (λt). So

$$P(D(t) = n) = P(N(t+s) - N(s) = n) = P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \text{ for } n = 0, 1, 2, ...$$

Also $E[D(t)] = V[D(t)] = E[N(t)] = V[N(t)] = \lambda t$.

58) Let X_1 be the waiting time until the 1st event, X_2 the waiting time from the 1st event until the 2nd event, ..., X_j the waiting time from the $j-1$th event until the jth event and so on. The X_i are called the waiting times or interarrival times. Let $S_n = \sum_{i=1}^n X_i$ the time of occurrence of the nth event = waiting time until the nth event. For a Poisson process with rate λ, the X_i are iid EXP(λ) with $E(X_i) = 1/\lambda$, and $S_n \sim$ Gamma (n, λ) with $E(S_n) = n/\lambda$ and $V(S_n) = n/\lambda^2$. Note that $S_n = S_{n-1} + X_n$ is a random walk with $S_n = Y_n$, $Y_0 = y_0 = 0$ and the $e_i = X_i \sim EXP(\lambda)$.

59) If the waiting times = interarrival times are iid EXP(λ), then $N(t)$ is a Poisson

process with rate λ.

60) Suppose $N(t)$ is a Poisson process with rate λ that counts events of k distinct types where $p_i = P($ type i event$)$. If $N_i(t)$ counts type i events, then $N_i(t)$ is a Poisson process with rate $\lambda_i = \lambda p_i$, and the $N_i(t)$ are independent for $i = 1, ..., k$. Then $N(t) = \sum_{i=1}^k N_i(t)$ and $\lambda = \sum_{i=1}^k \lambda_i$ where $\sum_{i=1}^k p_i = 1$.

61) A counting process $\{N(t) : t \geq 0\}$ is a *nonhomogeneous Poisson process* with *intensity function* or *rate function* $\lambda(t)$, also called a *nonstationary Poisson process*, and has the following properties. i) $N(0) = 0$. ii) The process has independent increments.

iii) $N(t)$ is a Poisson $m(t)$ RV where $m(t) = \int_0^t \lambda(r)dr$, and $N(t)$ counts the number of events that occurred in $(0, t]$ (or $(0, t)$).

iv) Let $0 < t_1 < t_2$. The RV $N(t_2) - N(t_1) \sim$ Poisson $(m(t_2) - m(t_1))$ where $m(t_2) - m(t_1) = \int_{t_1}^{t_2} \lambda(r)dr$ and $N(t_2) - N(t_1)$ counts the number of events that occurred in $(t_1, t_2]$ or (t_1, t_2).

62) If $N(t)$ is a Poisson process with rate λ and there are k distinct events where the probability $p_i(s)$ of the ith event at time s depends s, let $N_i(t)$ count type i events. Then $N_i(t)$ is a nonhomogeneous Poisson process with $\lambda_i(t) = \lambda \int_0^t p_i(s)ds$. Here $\sum_{i=1}^k p_i(s) = 1$ and the $N_i(t)$ are independent for $i = 1, ..., k$.

63) A stochastic process $\{X(t) : t \geq 0\}$ is a *compound Poisson process* if $X(t) = \sum_{i=1}^{N(t)} Y_i$ where $\{N(t) : t \geq 0\}$ is a Poisson process with rate λ and $\{Y_n : n \geq 0\}$ is a family of iid random variables independent of $N(t)$. The parameters of the compound process are λ and $F_Y(y)$ where $E(Y_1)$ and $E(Y_1^2)$ are important. Then $E[X(t)] = \lambda t E(Y_1)$ and $V[X(t)] = \lambda t E(Y_1^2)$.

64) The compound Poisson process has independent and stationary increments. Fix $r, t > 0$. Then ${}_tX_r = X(r + t) - X(r)$ has the same distribution as the RV $X(t)$. Hence $E({}_tX_r) = \lambda t E(Y_1)$ and $V({}_tX_r) = \lambda t E(Y_1^2)$.

65) Let $M_Y(t)$ be the moment generating function (mgf) of Y_1. Then the mgf of the RV $X(t)$ is

$$M_{X(t)}(r) = \exp(\lambda t[M_Y(r) - 1]).$$