Math 483 Exam 3 is Thursday, April. 3. Sections 1.3, 3.3, 3.9, 4.2, 4.3, 4.9, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, part of 6.3, 6.4, 6.5, part of 7.2, and 7.3 will be emphasized. 9 sheets of notes. **From exams 1 and 2:** $E(c) = c$, $E(cg(Y)) = cE(g(Y))$, and $E[\sum_{i=1}^k g_i(Y)] = \sum_{i=1}^k E[g_i(Y)]$ where c is any constant. **Six pages of notes but consider using the last 2 pages of the Exam 2 review as two of the six pages.**

1) **COMMON FINAL PROBLEM.** Given the probability function $p(y)$ if Y is discrete (eg from a table) 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. E1 2, E2 1?, Q2 5, Q4 4b, HW5 C), D), HW7 A), B), HW9 A)a, C), D).

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 $\int_{-\infty}^{\infty} f(t)dt = 1$.

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

2) **Common Problem.** Given that $f(y) = c\ g(y)$, find c. E2 4a?, Q4 4a, HW6 I), HW7 E).

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

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

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 $V(Z) = 1$.

If Y is **normal** $N(\mu, \sigma^2)$, then the pdf of Y is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)$$

where $\sigma > 0$ and μ and y are real. $E(Y) = \mu$ and $V(Y) = \sigma^2$.

You need to know how to use the standard normal table for exams 3, 4 and the final.

If Y is **exponential** (β) then the pdf of Y is $f(y) = \dfrac{1}{\beta}\exp(-\dfrac{y}{\beta}), y \geq 0$ where $\beta > 0$. The distribution function of Y is $F(y) = 1 - \exp(-y/\beta),\ y \geq 0$. $E(Y) = \beta$, and $V(Y) = \beta^2$. If W is exponential (β) then W is gamma $(\alpha = 1, \beta)$.

Integration. Know how to do u-substitution (and integration by parts).

Let Y_1 and Y_2 be discrete random variables. Then the **joint probability 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 probability function 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 distribution function** of any two random variables Y_1 and Y_2 is $F(y_1, y_2) = P(Y_1 \le y_1, Y_2 \le y_2), \forall y_1, y_2$.

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 $f(y_1, y_2) \ge 0, \forall y_1, y_2$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1$.

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

$F(y_1, ..., y_n) = P(Y_1 \le y_1, ..., Y_n \le y_n)$. 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 $F(y_1, ..., y_n) = \int_{-\infty}^{y_n} \cdots \int_{-\infty}^{y_1} f(t_1, ..., t_n) dt_1 \cdots dt_n$.

The **support** of a continuous RV is $\{y : f(y) > 0\}$. The support of jointly continuous (Y_1, Y_2) is $\{(y_1, y_2) : f(y_1, y_2) > 0\}$.

The **support** of a discrete RV is $\{y : p(y) > 0\}$. The support of jointly discrete (Y_1, Y_2) is $\{(y_1, y_2) : p(y_1, y_2) > 0\}$.

The *support* of the conditional probability function or 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 $0 \le y_1 \le y_2$.

Material after exam 2.

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) $f(y_1, y_2, ..., y_n) = f_{Y_1}(y_1) f_{Y_2}(y_2) \cdots f_{Y_n}(y_n)$.

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 $p(y_1, y_2) \ne p_{Y_1}(y_1) p_{Y_2}(y_2)$ 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.*

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

Q6 1a?, 2a?, HW11 A), B), C), D), E).

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.

4) **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 $Cov(Y_1, Y_2)$. E2 7?, Q6 1?, HW10 A)a, B), HW12 A), C).

5) **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. See E2 6?, Q5 3, 4, Q6 2?, HW10 C), D), E), HW12 B), D), E).

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 $Cov(U_1, U_2) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(Y_i, X_j)$.

$Y_1, ..., Y_n$ are a **random sample** or **iid** if they are independent and identically distributed (all come from the same population).

A **statistic** is a function of the random sample and known constants. The **sampling distribution** of a statistic is the population of the statistic.

Know: Let $Y_1, ..., Y_n$ be iid RV's with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$, then the **sample mean** $\overline{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$, $E(\overline{Y}) = \mu$ and $V(\overline{Y}) = \sigma^2/n$.

Know: If $Y_1, ..., Y_n$ be iid $N(\mu, \sigma^2)$, then $\overline{Y}$ is normal with mean $\mu_{\overline{Y}} = \mu$ and variance $\sigma^2_{\overline{Y}} = \sigma^2/n$.

The Central Limit Theorem (CLT): Let $Y_1, ..., Y_n$ be iid RV's with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$. Then $U_n = \sqrt{n}(\frac{\overline{Y} - \mu}{\sigma})$ converges in distribution to the standard normal distribution as $n \to \infty$.

Note that $U_n = \frac{\overline{Y} - E(\overline{Y})}{SD(\overline{Y})} = \frac{\overline{Y} - \mu_{\overline{Y}}}{\sigma_{\overline{Y}}} = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}$ is the z-score for $\overline{Y}$ where $\sigma_{\overline{Y}} = \sigma/\sqrt{n}$.

6) **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.
See Q7 3?, HW 15 C), D), E).

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.

7) **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. See Q7, 2?.

8) **Common Problem.** Let $Y_1, ..., Y_n$ be independent RV's with $E(Y_i) = \mu_i$, $V(Y_i) = \sigma_i^2$, and mgf $m_{Y_i}(t)$. Let $U = \sum_{i=1}^n Y_i$. Then find
a) $E(U) = E(\sum_{i=1}^n Y_i) = \sum_{i=1}^n E(Y_i) = \sum_{i=1}^n \mu_i$,
b) $V(U) = V(\sum_{i=1}^n Y_i) = \sum_{i=1}^n V(Y_i) = \sum_{i=1}^n \sigma_i^2$, and
c) the moment generating function (mgf) of U is $m_U(t) = \prod_{i=1}^n m_{Y_i}(t) = m_{Y_1}(t) m_{Y_2}(t) \cdots m_{Y_n}(t)$.
See Q7 5?, HW14 F), G), H).

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, $\prod_{i=1}^n \exp(b\theta_i) = \exp(\sum_{i=1}^n b\theta_i)$.

9) **Common problem (Not in Text):** Suppose that Y is a discrete RV with probability function $p_Y(y)$ given by a table. Let the **transformation** $U = h(Y)$ for some function h and find the probability function $p_U(u)$.
Solution: Step 1) Find $h(y)$ for each value of y.
Step 2) Collect $y : h(y) = u$, sum the corresponding probabilities:
$p_U(u) = \sum_{y:h(y)=u} p_Y(y)$, and table the result.
For example, if $U = Y^2$ and $p_Y(-1) = 1/3, p_Y(0) = 1/3$, and $p_Y(1) = 1/3$, then $p_U(0) = 1/3$ and $p_U(1) = 2/3$.
See Q7 4?.

Suppose that Y is a continuous RV with pdf $f_Y(y)$ on support $\mathcal{Y}$. Let the transformation $U = h(Y)$ for some function h and find the pdf $f_U(u)$ and the support $\mathcal{U}$ of U. This can be done with **the method of distribution functions or the method of transformations**.

Know: There are two ways to find the support $\mathcal{U}$ of $U = h(Y)$ if the support of Y is $\mathcal{Y} = [a, b]$. First, just plug in $h(y)$ and find the minimum and maximum value on $[a, b]$. A graph can help. If h is an increasing function, then $\mathcal{U} = [h(a), h(b)]$. If h is an decreasing function, then $\mathcal{U} = [h(b), h(a)]$. The second method is to find $y = h^{-1}(u)$. Then solve $a \le h^{-1}(u) \le b$ in terms of u.

The method of distribution functions: Suppose that the distribution function $F_Y(y)$ is known and that $U = h(Y)$.
a) If h is an increasing function then, $F_U(u) = P(U \le u) = P(h(Y) \le u) =$
$P(Y \le h^{-1}(u)) = F_Y(h^{-1}(u))$.
b) If h is a decreasing function then, $F_U(u) = P(U \le u) = P(h(Y) \le u) =$
$P(Y \ge h^{-1}(u)) = 1 - P(Y < h^{-1}(u)) = 1 - P(Y \le h^{-1}(u)) = 1 - F_Y(h^{-1}(u))$.
c) The special case $U = Y^2$ is important. If the support of Y is positive, use a). If the support of Y is negative, use b). If the support of Y is $[-a, a]$ (where $a = \infty$ is allowed), then $F_U(u) = P(U \le u) =$
$P(Y^2 \le u) = P(-\sqrt{u} \le Y \le \sqrt{u}) =$

$$\int_{-\sqrt{u}}^{\sqrt{u}} f_Y(y)dy = F_Y(\sqrt{u}) - F_Y(-\sqrt{u}),\ 0 \le u \le a^2.$$

After finding the distribution function $F_U(u)$, the pdf of U is $f_U(u) = \frac{d}{du}F_U(u)$ for $u \in \mathcal{U}$. Often the chain rule is needed. For example if the support of Y is $[-a, a]$ and if $U = Y^2$, then

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

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

The method of transformations: Assume that Y has pdf $f_Y(y)$ and support $\mathcal{Y}$. If $h(y)$ is either increasing or decreasing on $\mathcal{Y}$, then $U = h(Y)$ has pdf

$$f_U(u) = f_Y(h^{-1}(u))\ |\frac{dh^{-1}(u)}{du}|$$

on the support $\mathcal{U}$ of U. To be useful, this formula should be simplified as much as possible.

Tips: a) The pdf of U will often be that of a gamma RV. In particular, the pdf of U is often the pdf of an exponential(β) RV.
b) To find the inverse function $y = h^{-1}(u)$, solve the equation $u = h(y)$ for y.
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.

10) **COMMON FINAL PROBLEM:** Suppose that Y is a continuous RV with pdf $f_Y(y)$ on support $\mathcal{Y}$. Let the transformation $U = h(Y)$ for some function h and find the pdf $f_U(u)$ and the support $\mathcal{U}$ of U. To find the support of U, try evaluating $h(y)$ at the smallest and largest values from the support of Y. Usually you will use the method of transformations and h will be a monotone function, but sometimes the method of distribution functions will be used. See Q7 1?, HW13 C), HW 14 A), B), C), D), HW16 E).

11) **Common Problem:** Given two independent RV's X and Y, find $E(aX \pm bY) = aE(X) \pm bE(Y)$ and $V(aX \pm bY) = a^2V(X) + b^2V(Y)$. See HW12 A)c, B)c, HW13 B), C)e.