Math 581 Exam 2 is Thursday, Oct. 21, 2:00-3:15 NO NOTES. CHECK FORMULAS: YOU ARE RESPONSIBLE FOR ANY ERRORS ON THIS HANDOUT!

37) Fix $(\Omega, \mathcal{F}, P)$. A *simple random variable* (SRV) is a function $X : \Omega \to \mathbb{R}$ such that the range of X is finite and $\{X = x\} = \{\omega : X(\omega) = x\} \in \mathcal{F} \ \forall x \in \mathbb{R}$. Hence X is a discrete RV with finite support. Note that $X = \sum_{i=1}^n x_i I_{A_i}$ is a SRV if each $A_i \in \mathcal{F}$.

38) Suppose events $A_1, ..., A_n$ are disjoint and $\bigcup_{i=1}^n A_i = \Omega$. Let $X = \sum_{i=1}^n x_i I_{A_i}$. Then the expected value of X is $E(X) = \displaystyle\sum_{i=1}^n x_i P(A_i) = \sum_x x P(X = x)$ which is a finite sum since X is a SRV. The middle term is useful for proofs. For this SRV, $E(X)$ exists and is unique. In the second sum, the x need to be the distinct values in the range of X.

39) Suppose SRV X takes on distinct values $x_1, ..., x_m$. Then $X = \sum_{i=1}^m x_i I_{B_i}$ where the $B_i = \{X = x_i\}$ are disjoint with $\bigcup_{i=1}^n B_i = \Omega$. Hence a SRV has the form of 38) with $A_i = B_i$ and $n = m$.

40) Th. Let X_n, X and Y be SRVs.

a) $-\infty < E(X) < \infty$

b) linearity: $E(aX + bY) = aE(X) + bE(Y)$

c) If $X \le Y$, then $E(X) \le E(Y)$

d) If $\{X_n\}$ is uniformly bounded and $X = \lim_n X_n$ on a set of probability 1, then $E(X) = \lim_n E(X_n)$.

e) If t is a real valued function, then $E[t(X)] = \sum_x t(x) P(X = x)$

f) If X is nonnegative, $X \ge 0$, then $E(X) = \sum_i P(X > x_i) = \int_0^\infty [1 - F(x)]dx$.

41) **For the theory of integration**, assume the function f in the integrand is measurable where $f : \Omega \to \mathbb{R}$ and $(\Omega, \mathcal{F}, \mu)$ is a measure space.

42) A function $f : \Omega \to [-\infty, \infty]$ is a *measurable function* (or measurable or $\mathcal{F}$ measurable or Borel measurable) if

i) $f^{-1}(B) \in \mathcal{F} \ \forall B \in \mathcal{B}(\mathbb{R})$,

ii) $f^{-1}(\{\infty\}) = \{\omega : f(\omega = \infty\} \in \mathcal{F}$, and

iii) $f^{-1}(\{-\infty\}) = \{\omega : f(\omega = -\infty\} \in \mathcal{F}$.

43) Def. Let $f : \Omega \to [0, \infty]$ be a nonnegative function. Then the **integral**

$\displaystyle\int f d\mu = sup_{\{A_i\}} \sum_i (inf_{\omega \in A_i} f(\omega)) \mu(A_i)$ where $\{A_i\}$ is a finite $\mathcal{F}$ decomposition.)

(A finite $\mathcal{F}$ decomposition ($\mathcal{F}$ decomp of Ω) means that $A_i \in \mathcal{F}$ and $\Omega = \bigcup_{i=1}^n A_i$ for some n, and the A_i are disjoint.

44) Conventions for integration of a nonnegative function. a) $A_i = \emptyset$ implies that the inf term $= \infty$, b) $x(\infty) = \infty$ for $x > 0$, and c) $0(\infty) = 0$.

45) Theorem: Let $f \ge 0$ with $f(\omega) = \sum_{j=1}^m x_j I_{B_j}(\omega)$where each $x_j \ge 0$ and $\{B_j\}$ is an $\mathcal{F}$ decomp of Ω. Then $\int f d\mu = \sum_{j=1}^m x_j \mu(B_j)$.

46) If $f : \Omega \to [-\infty, \infty]$, then the **positive part** $f^+ = f I(f \ge 0) = max(f, 0)$, and the **negative part** $f^- = -f I(f \le 0) = max(-f, 0) = -min(f, 0)$. Hence $f^+(\omega) = f(\omega) I(f(\omega) \ge 0)$ and $f^-(\omega) = -f(\omega) I(f(\omega) \le 0)$.

Here $I(f \ge 0) = I(0 \le f \le \infty)$ while $I(f(\omega) \le 0) = I(-\infty \le f \le 0)$. If f is measurable, then $f^+ \ge 0$, $f^- \ge 0$ are both measurable, $f = f^+ - f^-$, and $|f| = f^+ + f^-$.

47) Convention: $\infty - \infty = -\infty + \infty$ is undefined.

48) Def: Let $f : \Omega \to [-\infty, \infty]$.

i) The **integral** $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$.
ii) The **integral is defined** unless it involves $\infty - \infty$.
iii) The function f is **integrable** if both $\int f^+ d\mu$ and $\int f^- d\mu$ are finite. Thus $\int f d\mu \in \mathbb{R}$ if f is integrable.

49) A property holds **almost everywhere** (ae), if the property holds for ω outside a set of measure 0, i.e. the property holds on a set A such that $\mu(A^c) = 0$. If μ is a probability measure P, then $P(A) = 1$ while $P(A^c) = 0$.

50) Theorem: suppose f and g are both nonnegative.
i) If $f = 0$ ae, then $\int f d\mu = 0$.
ii) If $\mu(\{\omega : f(\omega) > 0\}) > 0$, then $\int f d\mu > 0$.
iii) If $\int f d\mu < \infty$, then $f < \infty$ ae.
iv) If $f \le g$ ae, then $\int f d\mu \le \int g d\mu$.
v) If $f = g$ ae, then $\int f d\mu = \int g d\mu$.

51) Theorem: i) f is integrable iff $\int |f| d\mu < \infty$.
ii) **monotonicity**: If f and g are integrable and $f \le g$ ae, then $\int f d\mu \le \int g d\mu$.
iii) **linearity**: If f and g are integrable and $a, b \in \mathbb{R}$, then $af + bg$ is integrable with $\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$.
iv) **Monotone Convergence Theorem** (MCT): If $0 \le f_n \uparrow f$ ae, then $\int f_n d\mu \uparrow \int f d\mu$.
v) **Fatou's Lemma**: For nonnegative f_n, $\int liminf_n \; f_n d\mu \le liminf_n \; \int f_n d\mu$.
vi) **Lebesgue's Dominated Convergence Theorem** (LDCT): If the $|f_n| \le g$ ae where g is integrable, and if $f_n \to f$ ae, then f and f_n are integrable and $\int f_n d\mu \to \int f d\mu$.
vii) **Bounded Convergence Theorem** (BCT): If $\mu(\Omega) < \infty$ and the f_n are uniformly bounded, then $f_n \to f$ ae implies $\int f_n d\mu \to \int f d\mu$.
viii) If $f_n \ge 0$ then $\int \sum_{n=1}^\infty f_n d\mu = \sum_{n=1}^\infty \int f_n d\mu$.
ix) If $\sum_{n=1}^\infty \int |f_n| d\mu < \infty$, then $\int \sum_{n=1}^\infty f_n d\mu = \sum_{n=1}^\infty \int f_n d\mu$.
x) If f and g are integrable, then $|\int f d\mu - \int g d\mu| \le \int |f - g| d\mu$.

52) Consequences: a) linearity implies $\int \sum_{n=1}^k f_n d\mu = \sum_{n=1}^k \int f_n d\mu$: i.e., the integral and finite sum operators can be interchanged
b) MCT, LDCT, and BCT give conditions where the limit and $\int$ can be interchanged: $\lim_n \int f_n d\mu = \int \lim_n f_n d\mu = \int f d\mu$
c) 51) viii) and ix) give conditions where the infinite sum $\sum_{n=1}^\infty$ and the integral $\int$ can be interchanged: $\int \sum_{n=1}^\infty f_n d\mu = \sum_{n=1}^\infty \int f_n d\mu$.

53) A common technique is to show the result is true for indicators. Extend to simple functions by linearity, and then to nonnegative function by a monotone passage of the limit. Use $f = f^+ - f^-$ for general functions.

54) Induction Theorem: If $R(n)$ is a statement for each $n \in \mathbb{N}$ such that a) R(1) is true, and b) for each $k \in \mathbb{N}$, if $R(k)$ is true, then $R(k+1)$ is true,
then $R(n)$ is true for each $n \in \mathbb{N}$.

Note that $\infty \notin \mathbb{N}$. Induction can be used with linearity to prove 52) a), but induction generally does not work for 52) c).

55) Def. If $A \in \mathcal{F}$, then $\int_A f d\mu = \int f I_A d\mu$.

56) If $\mu(A) = 0$, then $\int_A f d\mu = 0$.

57) If $\mu : \mathcal{F} \to [0, \infty]$ is a measure and $f \ge 0$, then

a) $\nu(A) = \int_A f d\mu$ is a measure on $\mathcal{F}$.
b) If $\int_\Omega f d\mu = 1$, then $P(A) = \int_A f d\mu$ is a probability measure on $\mathcal{F}$.

58) For expected values, assume $(\Omega, \mathcal{F}, P)$ is fixed, and the random variables are measurable wrt $\mathcal{F}$.

59) We can define the expected value to be $E(X) = \int X dP$ as the special case of integration where the measure $\mu = P$ is a probability measure, or we can use a definition that ignores most measure theory.

60) Def. Let $X \geq 0$ be a nonnegative RV.
a) $E(X) = lim_{n\to\infty} E(X_n) = \displaystyle\int X dP \leq \infty$ where the X_n are nonnegative SRVs with $0 \leq X_n \uparrow X$.
b) The expectation of X over an event A is $E(XI_A)$.

There are several equivalent ways to define integrals and expected values. Hence $E(X)$ can also be defined as in 43) with μ replaced by P and f replaced by $X : \Omega \to \mathbb{R}$.

61) Theorem: Let X, Y be nonnegative random variables.
a) For $X, Y \geq 0$ and $a, b \geq 0$, $E(aX + bY) = aE(X) + bE(Y)$.
b) If $X \leq Y$ ae, then $E(X) \leq E(Y)$.

By induction, if the $a_i X_i \geq 0$, then $E(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n E(a_i X_i)$: the expected value of a finite sum of nonnegative RVs is the sum of the expected values.

62) For a random variable $X : \Omega \to (-\infty, \infty)$, then the **positive part** $X^+ = XI(X \geq 0) = max(X, 0)$, and the **negative part** $X^- = -XI(X \leq 0) = max(-X, 0) = -min(X, 0)$. Hence $X = X^+ - X^-$, and $|X| = X^+ + X^-$. Random variables are real functions: $\pm\infty$ are not allowed.

63) Def: Let the random variable $X : \Omega \to (-\infty, \infty)$.
i) The **expected value** $E(X) = \int X dP = \int X^+ dP - \int X^- dP = E(X^+) - E(X^-)$.
ii) The **expected value is defined** unless it involves $\infty - \infty$.
iii) The random variable X is **integrable** if $E[|X|] < \infty$. Thus $E(X) \in \mathbb{R}$ if X is integrable.

64) Theorem: i) X is integrable iff both $E[X^+]$ and $E[X^-]$ are finite.
ii) **monotonicity**: If X and Y are integrable and $X \leq Y$ ae, then $E(X) \leq E(Y)$.
iii) **linearity**: If X and Y are integrable and $a, b \in \mathbb{R}$, then $aX + bY$ is integrable with $E(aX + bY) = aE(X) + bE(Y)$.
iv) **Monotone Convergence Theorem** (MCT): If $0 \leq X_n \uparrow X$ ae, then $E(X_n) \uparrow E(X)$.
v) **Fatou's Lemma**: For RVs $X_n \geq 0$, $E[\liminf_n X_n] \leq \liminf_n E[X_n]$.
vi) **Lebesgue's Dominated Convergence Theorem** (LDCT): If the $|X_n| \leq Y$ ae where Y is integrable, and if $X_n \to X$ ae, then X and X_n are integrable and $E(X_n) \to E(X)$.
vii) **Bounded Convergence Theorem** (BCT): If the X_n are uniformly bounded, then $X_n \to X$ ae implies $E(X_n) \to E(X)$.
viii) If $X_n \geq 0$ then $E(\sum_{n=1}^\infty X_n) = \sum_{n=1}^\infty E(X_n)$.
ix) If $\sum_{n=1}^\infty E(|X_n|) < \infty$, then $E(\sum_{n=1}^\infty X_n) = \sum_{n=1}^\infty E(X_n)$.
x) If X and Y are integrable, then $|E(X) - E(Y)| \leq E[|X - Y|]$.

65) Consequences: a) linearity implies $E(\sum_{n=1}^k a_n X_n) = \sum_{n=1}^k a_n E(X_n)$: i.e., the

expectation and finite sum operators can be interchanged, or the expectation of a finite sum is the sum of the expectations if the X_n are integrable.
b) MCT, LDCT, and BCT give conditions where the limit and E can be interchanged: $\lim_n E(X_n) = E[\lim_n X_n] = E(X)$
c) 64) viii) and ix) give conditions where the infinite sum $\sum_{n=1}^{\infty}$ and the expected value can be interchanged: $E[\sum_{n=1}^{\infty} X_n] = \sum_{n=1}^{\infty} E(X_n)$.

66) Given $(\Omega, \mathcal{F}, P)$, the collection of all integrable random vectors or RVs is denoted by $L^1 = L^1(\Omega, \mathcal{F}, P)$.

67) Let $\boldsymbol{X}$ be a $1 \times k$ random vector with cdf $F_{\boldsymbol{X}}(\boldsymbol{t}) = F(\boldsymbol{t}) = P(X_1 \le t_1, ..., X_k \le t_k)$. Then the Lebesgue Stieltjes integral $E[h(\boldsymbol{X})] = \int h(\boldsymbol{t}) dF(\boldsymbol{t})$ provided the expected value exists, and the integral is a linear operator with respect to both h and F. If X is a random variable, then $E[h(X)] = \int h(t) dF(t)$. If $W = h(X)$ is integrable or if $W = h(X) \ge 0$, then the expected value exists. Here $h : \mathbb{R}^k \to \mathbb{R}^j$ with $1 \le j \le k$.

68) The distribution of a $1 \times k$ random vector $\boldsymbol{X}$ is a **mixture distribution** if the cdf of $\boldsymbol{X}$ is

$$F_{\boldsymbol{X}}(\boldsymbol{t}) = \sum_{j=1}^{J} \pi_j F_{\boldsymbol{U}_j}(\boldsymbol{t})$$

where the probabilities π_j satisfy $0 \le \pi_j \le 1$ and $\sum_{j=1}^{J} \pi_j = 1$, $J \ge 2$, and $F_{\boldsymbol{U}_j}(\boldsymbol{t})$ is the cdf of a $1 \times k$ random vector $\boldsymbol{U}_j$. Then $\boldsymbol{X}$ has a mixture distribution of the $\boldsymbol{U}_j$ with probabilities π_j. If X is a random variable, then

$$F_X(t) = \sum_{j=1}^{J} \pi_j F_{U_j}(t).$$

69) **Expected Value Theorem**: Assume all expected values exist. Let $d\boldsymbol{x} = dx_1 dx_2 ... dx_k$. Let $\mathcal{X}$ be the support of $\boldsymbol{X} = \{\boldsymbol{x} : f(\boldsymbol{x}) > 0\}$ or $\{\boldsymbol{x} : p(\boldsymbol{x}) > 0\}$.
a) If $\boldsymbol{X}$ has (joint) pdf $f(\boldsymbol{x})$, then $E[h(\boldsymbol{X})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(\boldsymbol{x}) f(\boldsymbol{x}) \, d\boldsymbol{x} = \int \cdots \int_{\mathcal{X}} h(\boldsymbol{x}) f(\boldsymbol{x}) \, d\boldsymbol{x}$. Hence $E[\boldsymbol{X}] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \boldsymbol{x} f(\boldsymbol{x}) \, d\boldsymbol{x} = \int \cdots \int_{\mathcal{X}} \boldsymbol{x} f(\boldsymbol{x}) \, d\boldsymbol{x}$.
b) If X has pdf $f(x)$, then $E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) \, dx = \int_{\mathcal{X}} h(x) f(x) \, dx$. Hence $E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{\mathcal{X}} x f(x) \, dx$.
c) If $\boldsymbol{X}$ has (joint) pmf $p(\boldsymbol{x})$, then $E[h(\boldsymbol{X})] = \sum_{x_1} \cdots \sum_{x_k} h(\boldsymbol{x}) p(\boldsymbol{x}) = \sum_{\boldsymbol{x} \in \mathbb{R}^k} h(\boldsymbol{x}) p(\boldsymbol{x}) = \sum_{\boldsymbol{x} \in \mathcal{X}} h(\boldsymbol{x}) p(\boldsymbol{x})$. Hence $E[\boldsymbol{X}] = \sum_{x_1} \cdots \sum_{x_k} \boldsymbol{x} p(\boldsymbol{x}) = \sum_{\boldsymbol{x} \in \mathbb{R}^k} \boldsymbol{x} p(\boldsymbol{x}) = \sum_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{x} p(\boldsymbol{x})$.
d) If X has pmf $p(x)$, then $E[h(X)] = \sum_x h(x) p(x) = \sum_{x \in \mathcal{X}} h(x) p(x)$. Hence $E[X] = \sum_x x p(x) = \sum_{x \in \mathcal{X}} x p(x)$.
e) Suppose $\boldsymbol{X}$ has a mixture distribution given by 68) and that $E(h(\boldsymbol{X}))$ and the $E(h(\boldsymbol{U}_j))$ exist. Then

$$E[h(\boldsymbol{X})] = \sum_{j=1}^{J} \pi_j E[h(\boldsymbol{U}_j)] \text{ and } \mathrm{E}(\boldsymbol{X}) = \sum_{\mathrm{j}=1}^{\mathrm{J}} \pi_{\mathrm{j}} \mathrm{E}[\boldsymbol{U}_{\mathrm{j}}].$$

f) Suppose X has a mixture distribution given by 68) and that $E(h(X))$ and the $E(h(U_j))$ exist. Then

$$E[h(X)] = \sum_{j=1}^{J} \pi_j E[h(U_j)] \text{ and } \mathrm{E(X)} = \sum_{\mathrm{j}=1}^{\mathrm{J}} \pi_{\mathrm{j}} \mathrm{E}[\mathrm{U_j}].$$

This theorem is easy to prove if the $\boldsymbol{U}_j$ are continuous random vectors with (joint) probability density functions (pdfs) $f_{\boldsymbol{U}_j}(\boldsymbol{t})$. Then $\boldsymbol{X}$ is a continuous random vector with pdf

$$f_{\boldsymbol{X}}(\boldsymbol{t}) = \sum_{j=1}^{J} \pi_j f_{\boldsymbol{U}_j}(\boldsymbol{t}), \quad \text{and} \quad \mathrm{E}[\mathrm{h}(\boldsymbol{X})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{h}(\boldsymbol{t}) \mathrm{f}_{\boldsymbol{X}}(\boldsymbol{t}) \mathrm{d}\boldsymbol{t}$$

$$= \sum_{j=1}^{J} \pi_j \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(\boldsymbol{t}) f_{\boldsymbol{U}_j}(\boldsymbol{t}) d\boldsymbol{t} = \sum_{j=1}^{J} \pi_j E[h(\boldsymbol{U}_j)]$$

where $E[h(\boldsymbol{U}_j)]$ is the expectation with respect to the random vector $\boldsymbol{U}_j$.

Alternatively, with respect to a Lebesgue Stieltjes integral, $E[h(\boldsymbol{X})] = \int h(\boldsymbol{t}) dF(\boldsymbol{t})$ provided the expected value exists, and the integral is a linear operator with respect to both h and F. Hence for a mixture distribution, $E[h(\boldsymbol{X})] = \int h(\boldsymbol{t}) dF(\boldsymbol{t}) =$

$$\int h(\boldsymbol{t}) \, d\left[\sum_{j=1}^{J} \pi_j F_{\boldsymbol{U}_j}(\boldsymbol{t})\right] = \sum_{j=1}^{J} \pi_j \int h(\boldsymbol{t}) dF_{\boldsymbol{U}_j}(\boldsymbol{t}) = \sum_{j=1}^{J} \pi_j E[h(\boldsymbol{U}_j)].$$

70) Fix $(\Omega, \mathcal{F}, P)$. Let the **induced probability** $P_X = P_F$ be $P_X(B) = P[X^{-1}(B)]$ for any $B \in \mathcal{B}(\mathbb{R})$. Then $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ is a probability space. If $\boldsymbol{X}$ is a $1 \times k$ random vector, then the **induced probability** $P_{\boldsymbol{X}} = P_F$ be $P_{\boldsymbol{X}}(B) = P[\boldsymbol{X}^{-1}(B)]$ for any $B \in \mathcal{B}(\mathbb{R}^k)$. Then $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P_{\boldsymbol{X}})$ is a probability space.

Then $E[h(\boldsymbol{X})] = \int h(\boldsymbol{X}) \, dP = \int h(\boldsymbol{x}) \, dF(\boldsymbol{x}) = E_F[h] = \int h \, dP_{\boldsymbol{X}}$. Then $E[h(X)] = \int h(X) \, dP = \int h(x) \, dF(x) = E_F[h] = \int h \, dP_X$. Here $W = h(X)$ is a RV wrt $(\Omega, \mathcal{F}, P)$, while $Z = h$ is a RV wrt $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$.

71) Let $X : \Omega \to \mathbb{R}$. Let $A, B, B_n \in \mathcal{B}(\mathbb{R})$.
i) If $A \subseteq B$, then $X^{-1}(A) \subseteq X^{-1}(B)$.
ii) $X^{-1}(\cup_n B_n) = \cup_n X^{-1}(B_n)$.
iii) $X^{-1}(\cap_n B_n) = \cap_n X^{-1}(B_n)$.
iv) If A and B are disjoint, then $X^{-1}(A)$ and $X^{-1}(B)$ are disjoint.
v) $X^{-1}(B^c) = [X^{-1}(B)]^c$.
(The unions and intersections in ii) and iii) can be finite, countable or uncountable.)

72) Theorem: Fix $(\Omega, \mathcal{F}, P)$. Let $X : \Omega \to \mathbb{R}$. X is a measurable function iff X is a RV iff any one of the following conditions holds.
i) $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F} \ \ \forall \ B \in \mathcal{B}(\mathbb{R})$.
ii) $X^{-1}((-\infty, t]) = \{X \le t\} = \{\omega \in \Omega : X(\omega) \le t\} \in \mathcal{F} \ \forall t \in \mathbb{R}$.
iii) $X^{-1}((-\infty, t)) = \{X < t\} = \{\omega \in \Omega : X(\omega) < t\} \in \mathcal{F} \ \forall t \in \mathbb{R}$.
iv) $X^{-1}([t, \infty)) = \{X \ge t\} = \{\omega \in \Omega : X(\omega) \ge t\} \in \mathcal{F} \ \forall t \in \mathbb{R}$.
v) $X^{-1}((t, \infty)) = \{X > t\} = \{\omega \in \Omega : X(\omega) > t\} \in \mathcal{F} \ \forall t \in \mathbb{R}$.

73) Theorem: Let X, Y, and X_i be RVs on $(\Omega, \mathcal{F}, P)$.
a) $aX + bY$ is a RV for any $a, b \in \mathbb{R}$. Hence $\sum_{i=1}^n X_i$ is a RV.
b) $\max(X, Y)$ is a RV. Hence $\max(X_1, ..., X_n)$ is a RV.
c) $\min(X, Y)$ is a RV. Hence $\min(X_1, ..., X_n)$ is a RV.
d) XY is a RV. Hence $X_1 \cdots X_n$ is a RV.
e) X/Y is a RV if $Y(\omega) \ne 0 \ \forall \ \omega \in \Omega$.

f) $sup_n\ X_n$ is a RV.
g) $inf_n\ X_n$ is a RV.
h) $limsup_n\ X_n$ is a RV.
i) $liminf_n\ X_n$ is a RV.
j) If $lim_n X_n = X$, then X is a RV.
k) If $\sum_{n=1}^{\infty} X_n \to X$, then X is a RV.
l) If $h : \mathbb{R}^n \to \mathbb{R}$ is measurable, then $Y = h(X_1, ..., X_n)$ is a RV.
m) If $h : \mathbb{R}^n \to \mathbb{R}$ is continuous, then h is measurable and $Y = h(X_1, ..., X_n)$ is a RV.
n) If $h : \mathbb{R} \to \mathbb{R}$ is monotone, then h is measurable and $h(X)$ is a RV.

74) Let $f(x) \geq 0$ be a Lebesgue integrable pdf of a RV with cdf F. Then $P_X(B) = P_F(B) = \int_B f(x)dx$ wrt Lebesgue integration. So many probability distributions can be obtained with Lebesgue integration.

75) RVs $X_1, ..., X_k$ are **independent** if $P(X_1 \in B_1, ..., X_k \in B_k) = \prod_{i=1}^n P(X_i \in B_i)$ for any $B_1, ..., B_k \in \mathcal{B}(\mathbb{R})$ iff $F_{X_1,...,X_k}(x_1, ..., x_k) = F_{X_1}(x_1) \cdots F_{X_k}(x_k)$ for any real $x_1, ..., x_k$ iff $\sigma(X_1), ..., \sigma(X_k)$ are independent ($\forall A_i \in \sigma(X_i)$, $A_1, ..., A_k$ are independent). An infinite collection of RVs $X_1, X_2, ...$ is **independent** if any finite subset is independent. If pdfs exist, $X_1, ..., X_k$ are independent iff $f_{X_1,...,X_k}(x_1, ..., x_k) = f_{X_1}(x_1) \cdots f_{X_k}(x_k)$ for any real $x_1, ..., x_k$. If pmfs exist, $X_1, ..., X_k$ are independent iff $p_{X_1,...,X_k}(x_1, ..., x_k) = p_{X_1}(x_1) \cdots p_{X_k}(x_k)$ for any real $x_1, ..., x_k$. Recall that the σ-field $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}\}$.

76) Suppose $X_1, ..., X_n$ are independent and $g_i(X_i)$ is a function of X_i alone. Then $E[g_1(x_1) \cdots g_n(X_i)] = E[\prod_{i=1}^n g_i(X_i)] = \prod_{i=1}^n E[g_i(X_i)]$ provided the expected values exist.

77) Let $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$ be two probability spaces. The **Cartesian product = cross product** $\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \Omega_1 \in \Omega_1, \Omega_2 \in \Omega_2\}$. The product of $\mathcal{F}_1$ and $\mathcal{F}_2$, denoted by $\mathcal{F}_1 \times \mathcal{F}_2$, is the σ-field $\sigma(\mathcal{A})$ where $\mathcal{A} = \{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$ is the collection of all cross products $A_1 \times A_2$ of events in $\mathcal{F}_1$ and $\mathcal{F}_2$.

78) Theorem: There is a unique probability measure $P = P_1 \times P_2$, called the product of P_1 and P_2 or the product probability measure, such that $P(A_1 \times A_2) = P_1(A_1)P_2(A_2)$ for all $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$.

79) The **product probability space** is $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, P_1 \times P_2)$.

80) 77)-79) can be extended to $(\Omega_i, \mathcal{F}_i, P_i)$ for $i = 1, ..., n$. Denote $P_1 \times \cdots \times P_n$ by $\prod_{i=1}^n P_i$, $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ by $\prod_{i=1}^n \mathcal{F}_i$, and $\Omega_1 \times \cdots \times \Omega_n$ by $\prod_{i=1}^n \Omega_i$. If $(\Omega_i, \mathcal{F}_i, P_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_i)$, then the product probability space is $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \prod_{i=1}^n P_i)$. If $(\Omega_i, \mathcal{F}_i, P_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{X_i})$, then the product probability space is $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \prod_{i=1}^n P_{X_i})$.

81) Let **independent** X_i be defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_{X_i})$. Then the product probability space $(\Omega, \mathcal{F}, P) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \prod_{i=1}^n P_{X_i})$ is the probability space for $\boldsymbol{X} = (X_1, ..., X_n)$.

82) Let $\int f d\mu = \int f(x)d\mu(x)$. Then the double integral

$$\int\int_{\Omega_1 \times \Omega_2} f(x_1, x_2) d[P_1 \times P_2(x_1, x_2)] =$$

$$\int_{\Omega_1} \left[\int_{\Omega_2} f(x_1, x_2) dP_2(x_2) \right] dP_1(x_1) = \int_{\Omega_2} \left[\int_{\Omega_1} f(x_1, x_2) dP_1(x_1) \right] dP_2(x_2).$$

The last two equations are known as iterated integrals.

83) **Fubini's Theorem:** a) Assume $f \geq 0$. Then $\int_{\Omega_1} f(x_1, x_2)dP_1(x_1)$ is measurable $\mathcal{F}_2$, $\int_{\Omega_2} f(x_1, x_2)dP_2(x_2)$ is measurable $\mathcal{F}_1$, and 82) holds.

b) Assume f is integrable wrt $P_1 \times P_2$, then $\int_{\Omega_1} f(x_1, x_2)dP_1(x_1)$ is finite ae and measurable $\mathcal{F}_2$ ae, $\int_{\Omega_2} f(x_1, x_2)dP_2(x_2)$ is finite ae and measurable $\mathcal{F}_1$ ae, and 82) holds.

Note: Part 83 a) is also known as Tonelli's theorem or the Fubini-Tonelli theorem. The double integral is often written as $\int_{\Omega_1 \times \Omega_2}$. Note that $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ (at least ae). Fubini's theorem for product probability measures shows double integrals can be calculated with iterated integrals if $X_1 \perp\!\!\!\perp X_2$, and the theorem is sometimes stated as below.

84) **Fubini's Theorem for product probability measures**: If f is measurable, then

$$\int_{\Omega_1 \times \Omega_2} f d[P_1 \times P_2] = \int_{\Omega_1} \left[\int_{\Omega_2} f(x_1, x_2)dP_2(x_2) \right] dP_1(x_1) = \int_{\Omega_2} \left[\int_{\Omega_1} f(x_1, x_2)dP_1(x_1) \right] dP_2(x_2)$$

provided that either a) $f \geq 0$, or b) $\int_{\Omega_1 \times \Omega_2} |f| d[P_1 \times P_2] < \infty$.

85) A **product measure** μ satisfies $\mu(\prod_{i=1}^n A_i) = \prod_{i=1}^n \mu(A_i)$.

86) **Fubini's Theorem for product measures**: If f is measurable, then

$$\int_{\Omega_1 \times \Omega_2} f d[\mu_1 \times \mu_2] = \int_{\Omega_1} \left[\int_{\Omega_2} f(x_1, x_2)d\mu_2(x_2) \right] d\mu_1(x_1) = \int_{\Omega_2} \left[\int_{\Omega_1} f(x_1, x_2)d\mu_1(x_1) \right] d\mu_2(x_2)$$

provided that the μ_i are σ-finite and either a) $f \geq 0$, or b) $\int_{\Omega_1 \times \Omega_2} |f| d[\mu_1 \times \mu_2] < \infty$.

Note: the Lebesgue measure is $\sigma-$finite on $\mathbb{R}$ and the counting measure μ_C is σ-finite if Ω is countable, where $\mu_C(A) =$ the number of points in set A. Let λ be the Legesgue measure on $\mathbb{R}^2$ and μ_L the Lebesgue measure on $\mathbb{R}$. The $\lambda(A \times B) = \mu_L(A)\mu_L(B)$ is a product measure. Let ν be the counting measure on $\mathbb{Z}^2$ and μ_C the counting measure on $\mathbb{Z}$. Then $\nu(A \times B) = \mu_C(A)\mu_C(B)$ is a product measure.

87) **Fubini's Theorem for Lebesgue Integrals**: Let $C = \{(x, y) : a \leq x \leq b, c \leq y \leq d\} = [a, b] \times [c, d]$. Let $g(x, y)$ be measurable and Lebesgue integrable. Then

$$\int \int_C g(x, y)dxdy = \int_c^d \left[\int_a^b g(x, y)dx \right] dy = \int_a^b \left[\int_c^d g(x, y)dy \right] dx.$$

88) The result in 87) can be extended to where the limits of integration are infinite and to $n \geq 2$ integrals. Using $g(x, y) = h(x, y)f(x, y)$ where f is a pdf gives $E[h(X, Y)]$. Note that $g : \mathbb{R}^2 \to \mathbb{R}$ (at least ae).

89) **Central Limit Theorem (CLT)**: Let $X_1, ..., X_n$ be iid with $E(X) = \mu$ and $V(X) = \sigma^2$. Then $\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{D} N(0, \sigma^2)$.

90) If F_n and F are cdfs, then F_n **converges weakly** to F, written $F_n \xrightarrow{W} F$, if $lim_n F_n(x) = F(x)$ at every continuity point of X.

91) Let $\{Z_n, n = 1, 2, ...\}$ be a sequence of random variables with cdfs F_n, and let X be a random variable with cdf F. Then Z_n **converges in distribution to** X, written

$$Z_n \xrightarrow{D} X,$$

or Z_n *converges in law to* X, written $Z_n \xrightarrow{L} X$, if

$$\lim_{n\to\infty} F_n(t) = F(t)$$

at each continuity point t of F. The distribution of X is called the **limiting distribution** or the **asymptotic distribution** of Z_n.

Notes: a) If $X_n \xrightarrow{D} X$, then the limiting distribution (the distribution of X) **does not** depend on n.

$$b) Z_n = \sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) = \left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}\right) = \left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}\right)$$

is the z–score of $\overline{X}_n$ (and the z-score of $\sum_{i=1}^n X_i$), and $Z_n \xrightarrow{D} N(0,1)$. c) Two applications of the CLT are to give the limiting distribution of $\sqrt{n}(\overline{X}_n - \mu)$ and the limiting distribution of $\sqrt{n}(X_n/n - \mu_Y)$ for a random variable X_n such that $X_n = \sum_{i=1}^n Y_i$ where the Y_i are iid with $E(Y) = \mu_Y$ and $V(Y) = \sigma_Y^2$. See point 92) below. d) $X_n \xrightarrow{D} X$ is equivalent to F_{X_n} converges weakly to F_X.

92) Theorem: a) If $Y_1, ..., Y_n$ are iid binomial BIN(k, ρ) random variables, then $X_n = \sum_{i=1}^n Y_i \sim$ BIN(nk, ρ). Note that $E(Y_i) = k\rho$ and $V(Y_i) = k\rho(1-\rho)$.

b) Denote a chi–square χ_p^2 random variable by $\chi^2(p)$. If $Y_1, ..., Y_n$ are iid χ_p^2, then $X_n = \sum_{i=1}^n Y_i \sim \chi_{np}^2$. Note that $E(Y_i) = p$ and $V(Y_i) = 2p$.

c) If $Y_1, ..., Y_n$ are iid exponential EXP$(\beta) \sim G(1, \beta)$, then $X_n = \sum_{i=1}^n Y_i \sim G(n, \beta)$. Note that $E(Y_i) = 1/\beta$ and $V(Y_i) = 1/\beta^2$.

d) If $Y_1, ..., Y_n$ are iid gamma $G(\alpha, \beta)$, then $X_n = \sum_{i=1}^n Y_i \sim G(n\alpha, \beta)$. Note that $E(Y_i) = \alpha/\beta$ and $V(Y_i) = \alpha/\beta^2$.

e) If $Y_1, ..., Y_n$ are iid $N(\mu, \sigma^2)$, then $X_n = \sum_{i=1}^n Y_i \sim N(n\mu, n\sigma^2)$. Note that $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$.

f) If $Y_1, ..., Y_n$ are iid Poisson $POIS(\theta)$, then $X_n = \sum_{i=1}^n Y_i \sim$ POIS(nθ). Note that $E(Y_i) = V(Y_i) = \theta$.

g) If $Y_1, ..., Y_n$ are iid inverse Gaussian $IG(\theta, \lambda)$, then $X_n = \sum_{i=1}^n Y_i \sim IG(n\theta, n^2\lambda)$. Note that $E(Y_i) = \theta$ and $V(Y_i) = \theta^3/\lambda$.

h) If $Y_1, ..., Y_n$ are iid geometric geom$(p) \sim$ NB(1, p), then $X_n = \sum_{i=1}^n Y_i \sim$ NB(n, p). Note that $E(Y_i) = (1-p)/p$ and $V(Y_i) = (1-p)/p^2$.

i) If $Y_1, ..., Y_n$ are iid negative binomial $NB(r, \rho)$, then $X_n = \sum_{i=1}^n Y_i \sim NB(nr, \rho)$. Note that $E(Y_i) = r(1-p)/p$ and $V(Y_i) = r(1-p)/p^2$.