If the events are not independent, then they are dependent.

p53 23] First Borel-Cantelli lemma: Let $(\Omega, \mathcal{F}, P)$ be fixed. If $\sum_{n=1}^{\infty} P(A_n)$ converges, then $\sum P(A_n) < \infty$

$$P(\limsup_n A_n) = 0. \quad (A_n \text{ events})$$

proof] Since $\limsup_n A_n \subseteq \bigcup_{k=m}^{\infty} A_k$ for any m,

$$P(\limsup_n A_n) \leq P\left(\bigcup_{k=m}^{\infty} A_k\right) \leq \sum_{k=m}^{\infty} P(A_k)$$

$\leq \varepsilon$ for $m \geq m(\varepsilon)$ by def of a convergent sum. Since $\varepsilon > 0$ is arbitrary, $P(\limsup_n A_n) = 0$. $\square$

p55 24] * Second Borel Cantelli lemma: Let $(\Omega, \mathcal{F}, P)$ be fixed. If $\{A_n\}$ is a sequence of independent events and $\sum_{n=1}^{\infty} P(A_n) = \infty$ (the sum diverges)

then $P(\limsup_n A_n) = 1$.

proof] Let $B_n = \bigcup_{k=n}^{\infty} A_k$. Then

$\limsup_n A_n \subseteq B_n$ and $B_n \perp \limsup_n A_n$.

and $\limsup_n A_n = \bigcap_{n=1}^{\infty} B_n$.

Thus $P(B_n) \downarrow P(\limsup_n A_n)$

and $\lim_n P(B_n) = P(\limsup_n A_n)$.

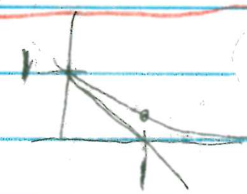
Claim: $P(B_n) = 1$, $n \geq 1$.

proof: $\underline{P\left(\bigcup_{k=n}^{N} A_k\right)} = 1 - P\left[\left(\bigcup_{k=n}^{N} A_k\right)^c\right]$

$= 1 - P\left(\bigcap_{k=n}^{N} A_k^c\right) \underset{\text{ind of } A_k}{=}$

$1 - \prod_{k=n}^{N} P(A_k^c) = \underline{1 - \prod_{k=n}^{N} [1 - P(A_k)]}$.

Fact: $1 - x \leq e^{-x}$, $0 \leq x \leq 1$



Thus $1 - P(A_k) \leq \exp[-P(A_k)]$.

$\therefore \prod_{k=n}^{N} [1 - P(A_k)] \leq \prod_{k=n}^{N} \exp[-P(A_k)]$

$= \exp\left[-\sum_{k=n}^{N} P(A_k)\right]$.

(*) So $1 - \prod_{k=n}^{N} [1 - P(A_k)] \geq 1 - \exp\left[-\sum_{k=n}^{N} P(A_k)\right]$

Thus $P\left(\bigcup_{k=n}^{N} A_k\right) \geq 1 - \exp\left[-\sum_{k=n}^{N} P(A_k)\right]$.

By hyp, $\sum_{k=n}^{\infty} P(A_k) = \infty$.

$e^{-\infty} = 0$. $\therefore$ Choose $N = N(\varepsilon)$
$\ni$ $\exp\left(-\sum_{k=n}^{N} P(A_k)\right) < \varepsilon$.

Then $P(B_n) = P\left(\bigcup_{k=n}^{\infty} A_k\right) \geq$
$P\left(\bigcup_{k=n}^{N} A_k\right) \geq 1-\varepsilon$.

ε was arbitrary, $\therefore$ $P(B_n) = 1$.

$\therefore$ $P\left(\bigcap_{n=1}^{\infty} B_n\right) = 1 = P(\limsup_n A_n)$.

$\square$

p46 25.) If $P(A_n) = 0$ $\forall n$, then

$0 \leq P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n) = 0$

So $P\left[\left(\bigcup_{n=1}^{\infty} A_n\right)^c\right] = 1-0 = 1 = P\left(\bigcap_{n=1}^{\infty} A_n^c\right)$

where $P(A_n^c) = 1$.
If $A_1, A_2, \ldots$ are sets of prob 0
so is $\bigcup_{n=1}^{\infty} A_n$. If $A_1, A_2, \ldots$ are
sets of prob 1 so is $\bigcap_{n=1}^{\infty} A_n$.

p57 25] Let $A_1, A_2, \ldots$ be a sequence of events in $(\Omega, \mathcal{F}, P)$. Let

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(A_n, A_{n+1}, \ldots)$$

be the <u>tail σ-field</u> associated with $\{A_n\}_{n=1}^{\infty}$. If $A \in \mathcal{T}$, then A is a <u>tail event</u>.

Note: $\sigma(A_n, A_{n+1}, \ldots)$ is the σ-field generated by $\mathcal{A}_n = \{A_k\}_{k=n}^{\infty}$.

p57-58 26] * Kolmogorov's 0-1 law:

Let $A_1, A_2, \ldots$ be a sequence of ind events in $(\Omega, \mathcal{F}, P)$. If $A \in \mathcal{T}$, then $P(A) = 0$ or $P(A) = 1$.

Key fact: If events $A_1, A_2, \ldots$ are ind, then $\sigma(A_{i_1}, \ldots, A_{i_n}) \perp\!\!\!\perp \sigma(A_{i_{n+1}}, \ldots, A_{i_m})$ if $\{i_1, \ldots, i_n\} \cap \{i_{n+1}, \ldots, i_m\} = \phi$.
Here $m = \infty$ and $n = \infty$ are allowed.
(proof omitted but see p 50.)

proof of the 0-1 law: Let $A \in \mathcal{T}$
want to show $A, A_1, A_2, \ldots$ are ind
if $A, A_n \in \mathcal{T}$. To do this, need to show every finite subcollection is ind. It is enough to show
$A, A_1, \ldots, A_{n-1}$ are ind for $n \geq 2$
(since every finite subcollection will be a subset of these collections for some n).
$\sigma(A_1), \sigma(A_2), \ldots, \sigma(A_{n-1}), \sigma(A_n, A_{n+1}, \ldots)$

are n σ-fields and the As have disjoint indices. Since $A \in \sigma(A_n, A_{n+1}, \ldots)$

$A, A_1, \ldots, A_{n-1}$ are ind.

$\therefore A, A_1, A_2, \ldots$ are ind.

$\therefore \sigma(A)$ and $\sigma(A_1, A_2, \ldots)$ are ind.
$A \in \sigma(A)$ and $A \in \mathcal{T} \subseteq \sigma(A_1, A_2, \ldots)$.

10.5

$\therefore A \perp\!\!\!\perp A$ and

$$P(A) = P(A \wedge A) \overset{\text{ind}}{=} P(A)\,P(A)$$

or $P(A)\,(1 - P(A)) = 0$. $\therefore$

$P(A) = 1$ or $P(A) = 0$.

(Note: $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$ are ind if for each choice of $A_i \in \mathcal{D}_i$, the events $A_1, \ldots, A_n$ are ind where the $\mathcal{D}_i$ are classes of $\mathcal{F}$ sets.)

§5, 13, 20 1) * Def: Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function $X : \Omega \to \mathbb{R} = (-\infty, \infty)$ is a **random variable** if the **inverse image**

$$X^{-1}(B) \in \mathcal{F} \quad \forall B \in \overbrace{\mathcal{B}(\mathbb{R})}^{\text{Borel }\sigma\text{-field on }\mathbb{R}},$$

Equivalently, $\{X \le t\} = \{\omega \in \Omega : X(\omega) \le t\} \in \mathcal{F} \quad \forall t \in \mathbb{R}$.

2) In 1), $\{X \le t\}$ is an event.

pm 11

$X^{-1}(B)$ is a set known as the inverse image. $X^{-1}(B)$ is not the inverse function.

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}.$$

$$X^{-1}\left(\bigcup_{\lambda} B_\lambda\right) = \bigcup_{\lambda} X^{-1}(B_\lambda)$$

$$X^{-1}\left(\bigcap_{\lambda} B_\lambda\right) = \bigcap_{\lambda} X^{-1}(B_\lambda).$$

where $\{B_\lambda\}$ is a collection of subsets of $\mathcal{B}(\mathbb{R})$.

The RV X is a measurable function.

Note: it will take a while to define the terms in 1 and 2.

p63 3) Let $(\Omega, \mathcal{F}, P)$ be a prob space.

X is simple random variable if the range of X is $\{x_1, \ldots, x_n\}$

and if $\{\omega : X(\omega) = x\} \in \mathcal{F}$

$\forall x \in \mathbb{R}$,

$\{\omega : X(\omega) = x\} = \phi \in \mathcal{F}$ if

$x \notin \{x_1, \ldots, x_n\}$.

$\{\omega : X(\omega) = x_i\} = X^{-1}(\{x_i\})$

for $x_i \in \{x_1, \ldots, x_n\}$.

4) For now, take the expected value of $X = E(X)$ and the variance of $X = V(X) = VAR(X)$ $= E[(X - E(X))^2] = E[X^2] - [E(X)]^2$ as in a calculus based course such as Math 483.

5)* Theorem: Let $u : \mathbb{R} \to [0, \infty)$ be a nonnegative function

i) Generalized Chebyshev's Inequality = Generalized Markov's Inequality: If $E[u(Y)]$ exists, then for any $c > 0$, $P(u(Y) \geq c) \leq \frac{E[u(Y)]}{c}$.

i) Markov's Inequality: Suppose $\mu = E(Y)$ exists. For any $r > 0$ and $c > 0$, $P(|Y-\mu| \geq c) =$

$$P[|Y-\mu|^r \geq c^r] \leq \frac{E[|Y-\mu|^r]}{c^r}.$$

p 75 ii) Know Chebyshev's inequality:

Suppose $VAR(Y)$ exists. Then for any $c > 0$, $P[|Y-\mu| \geq c] \leq \frac{VAR(Y)}{c^2}$.

Proof for probability density functions pdfs. For probability mass functions pmfs, replace integrals by sums.

i) $E[u(Y)] = \int_{\mathbb{R}} u(y) f(y)\, dy =$

$$\int_{\{y: u(y) \geq c\}} u(y) f(y)\, dy + \int_{\{y: u(y) < c\}} \underbrace{u(y) f(y)}_{\text{nonnegative}}\, dy$$

$$\geq \int_{\{y: u(y) \geq c\}} u(y) f(y)\, dy \geq c \int_{\{y: u(y) \geq c\}} f(y)\, dy$$

$$= c\, P[u(Y) \geq c].$$

ii) Take $U(Y) = |Y-\mu|^r$ and $\tilde{c} = c^r$.

Then $P[|Y-\mu| \geq c] = P[|Y-\mu|^r \geq c^r]$

$\leq \dfrac{E[|Y-\mu|^r]}{c^r}$ by i).

iii) Take $r = 2$. Then

$P[|Y-\mu| \geq c] = P[|Y-\mu|^2 \geq c^2]$

$\leq \dfrac{V(Y)}{c^2}$ by ii).

$\square$

p74 6) If $U(Y) = |Y|^k$ and $\tilde{c} = c^k$

$P[|Y| \geq c] = P[|Y|^k \geq c^k]$

$\leq \dfrac{E|Y|^k}{c^k}$.

Skip §6, §7, §8 for now.

p 142 §9. 6) The <u>moment generating function</u> (mgf) of a RV X

PM 13

is $M(t) = E[e^{tX}]$ if the expectation exists for t in some neighborhood of 0. Otherwise the mgf does not exist.

p144 7) If $M(t)$ exists, the <u>cumulant generating function</u> of X is $C(t) = \log M(t) = \log E[e^{tX}]$ where $\log = \log_e = \ln$ in this class

p142 5) Let $g^{(k)}(t)$ be the kth derivative of g with $g' = g^{(1)}$ and $g'' = g^{(2)}$.

i) Then $E(X^k) = M^{(k)}(0)$, the kth derivative of the mgf $M(t)$ evaluated at 0.

ii) $C'(0) = E(X)$, $C''(0) = V(X)$.

Skip the rest of §9 for now.

§2.10

p155 1] Let $\underline{x}=(x_1, \ldots, x_K) \in \mathbb{R}^K$.

Let $\mathcal{A}$ be the class of "rectangles" $\{\underline{x} \in \mathbb{R}^K : a_i < x_i \le b_i$, $i=1,\ldots,K\}$ where $a_i, b_i \in \mathbb{R}$.

$\mathcal{B}(\mathbb{R}^K) = \sigma(\mathcal{A})$ is the Borel σ-field on $\mathbb{R}^K$.

p157 2] μ is a measure on $(\Omega, \underset{\sigma\text{-field}}{\mathcal{F}})$

if m1) $\mu(A) \in [0, \infty]$ for $A \in \mathcal{F}$
∞ may be allowed

m2) $\mu(\phi) = 0$

m3) If $A_1, A_2, \ldots$ are disjoint, then $\mu\left(\bigcup_{K=1}^{\infty} A_K\right) = \sum_{K=1}^{\infty} \mu(A_K)$.
(countable additivity)

3] The measure μ is finite if $\mu(\Omega) < \infty$ and infinite if $\mu(\Omega) = \infty$.

If $\Omega = \bigcup_{i=1}^{\infty} A_i$, $A_i \in \mathcal{F}$

positive integers

and $\mu(A_k) < \infty$ for $k \in \mathbb{N}$,

then μ is <u>σ-finite.</u>

ex) If $\mu(\Omega) = 1$, then μ is

a probability measure and

μ is σ-finite.

4) $(\Omega, \mathcal{F}, \mu)$ is a <u>measure space</u>.

$(\Omega, \mathcal{F})$ is a <u>measurable space</u>.

5) properties of measures

$\infty \le \infty$

i) monotone: $\mu(A) \le \mu(B)$ if $A \subseteq B$ are $\mathcal{F}$ sets

ii) If $A_1, \ldots, A_n$ are disjoint $\mathcal{F}$ sets

then $\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k)$.

(finite additivity)

iii) If $A \subseteq B$ are $\mathcal{F}$ sets,

$\mu(B - A) = \mu(B) - \mu(A)$ if $\mu(B) < \infty$.

iv) Finite subadditivity: if $A_k \in \mathcal{F}$ then

14

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) \le \sum_{k=1}^{n} \mu(A_k)$$

and the A_k need not have finite measure.

v) continuity from below: if $A_n \in \mathcal{F}$ and $A_n \uparrow A$ then $\mu(A_n) \uparrow \mu(A)$.

vi) continuity from above: if $A_n \in \mathcal{F}$
$\mu(A_1)<\infty$ and $A_n \downarrow A$, then $\mu(A_n) \downarrow \mu(A)$.

vii) countable subadditivity:

If $A_k \in \mathcal{F}$ then $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \le \sum_{k=1}^{\infty} \mu(A_k)$.

Note: most of the proofs are similar to those for a prob measure P. For vi), since $\mu(A_1)<\infty$, $[A_1 - A_n] \uparrow [A_1 - A]$

$\Rightarrow \mu(A_1) - \mu(A_n) = \mu(A_1 - A_n) \uparrow \mu(A_1 - A) = \mu(A_1) - \mu(A)$. so $\mu(A_n) \downarrow \mu(A)$.

Skip §11, 12

§13 Measurable Functions and Mappings

1] Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be 2 measurable spaces. For a mapping $T: \Omega \to \Omega'$, the mapping T is measurable $\mathcal{F}/\mathcal{F}'$ if $T^{-1}(A') \in \mathcal{F}$ for each $A' \in \mathcal{F}'$.

2] $T^{-1}(A') = \{\omega \in \Omega : T(\omega) \in A'\}$ for $A' \subseteq \Omega'$.

3] * Let a real function $g: \Omega \to \mathbb{R}$, with $\Omega' = \mathbb{R}$ and $\mathcal{F}' = \mathcal{B}(\mathbb{R})$.

Then g is measurable or measurable $\mathcal{F}$ (measurable function or)

if $g^{-1}(B) = \{\omega : g(\omega) \in B\} \in \mathcal{F}$ for every $B \in \mathcal{B}(\mathbb{R})$.

Fact: g is measurable if $\{\omega : g(\omega) \le x\} \in \mathcal{F}$ $\forall x \in \mathbb{R}$.

4] * A random variable X is

15.5

a measurable function since

$X^{-1}(B) \in \mathcal{F} \quad \forall B \in \mathcal{B}(\mathbb{R})$.

5] $\therefore$ X is a random variable iff X is a measurable function.

§3, 10 Uniqueness

p.36 6] A class Π (pi) of subsets of Ω is a π-system if it is closed under the formation of finite intersections:
$A, B \in \Pi \Rightarrow A \cap B \in \Pi$.

p.3? 7] A class Λ (lambda) of subsets of Ω is a λ-system if it contains Ω and is closed under the formation of complements and of countable (including finite) disjoint unions.

a i) $\Omega \in \Lambda$

a ii) $A \in \Lambda \Rightarrow A^c \in \Lambda$

a iii) $A_1, A_2, \ldots \in \Lambda$ and

$A_i \cap A_j = \emptyset$ for $i \neq j \Rightarrow$

$\bigcup_n A_n \in \Lambda$.