University of New South Wales

School of Mathematics and Statistics

Assignment 2

Measure Theory

Author:
Adam J. Gray

Student Number:
3329798

1

1.1

Let μ and ν be probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We wish to show that the convolution

$$\mu * \nu(B) = \int \nu(B - x)\mu(dx) \tag{1}$$

of these two measures is well defined in the sense that $\nu(B - x)$ is measurable in x and the integral exists. Firstly see that

$$\nu(B - x) = \int \chi_{B-x}(y)\nu(dy) \tag{2}$$
$$= \int \chi_B(x + y)\nu(dy) \tag{3}$$

and by seeing that $\phi(x, y) = \chi_B(x+y)$ is $\mathcal{B}(R)\otimes\mathcal{B}(R)$ measurable and so applying Tonelli's theorem $\nu(B-x)$ is measurable.
We now need to show that the integral exists. This is clear because

$$0 \leq \nu(B - x) \leq 1 \tag{4}$$

and thus

$$\int \nu(B - x)\mu(dx) \leq \int 1\mu(dx) \tag{5}$$
$$= 1 \tag{6}$$

because μ is a probability measure.

1.2

Suppose there exists a bounded set $F \in \mathcal{B}(\mathbb{R})$ such that

$$\mu * \nu(F) = 1 \tag{7}$$

we wish to show that there exists bounded sets $G, H \in \mathcal{B}(\mathbb{R})$ such that

$$\mu(G) = 1 \quad \text{and} \quad \nu(H) = 1. \tag{8}$$

See that if

$$1 = \mu * \nu(F) \tag{9}$$
$$= \int \nu(F - x)\mu(dx) \tag{10}$$
$$= \int \int \chi_F(x + y)\nu(dy)\mu(dx) \tag{11}$$

Now as F is bounded there there exist intervals $G = [a, b]$ and $H = [c, d]$ such that

$$\int \int \chi_F(x + y)\chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \tag{12}$$

and further

$$1 = \int\int \chi_F(x+y)\chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \leq \int\int \chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \tag{13}$$
$$= \int \chi_H(y)\nu(dy)\int \chi_G(x)\mu(dx) \tag{14}$$

[1] which implies that

$$\int \chi_G(x)\mu(dx) = 1 \quad \text{and} \quad \int \chi_H(y)\nu(dy) = 1 \tag{15}$$

that is $\mu(G) = 1$ and $\nu(H) = 1$.
Suppose that F is now countable but such that $\mu * \nu(F) = 1$. Then

$$1 = \int \nu(F-x)\mu(dx) \tag{16}$$
$$= \int\int \chi_F(x+y)\nu(dy)\mu(dx) \tag{17}$$

now since $F = \{(x_k, y_k)\}_{k\in\mathbb{N}}$ is countable then there must exist countable sets $G = \{x_k\}_{k\in\mathbb{N}}$, $H = \{y_k\}_{k\in\mathbb{N}}$ such that

$$\int\int \chi_F(x+y)\nu(dy)\mu(dx) = \int\int \chi_F(x+y)\chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \tag{18}$$

and further

$$1 = \int\int \chi_F(x+y)\chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \leq \int\int \chi_G(x)\chi_H(y)\nu(dy)\mu(dx) \tag{19}$$
$$= \int \chi_H(y)\nu(dy)\int \chi_G(x)\mu(dx) \tag{20}$$

which as before implies $\mu(G) = 1$ and $\nu(H) = 1$.
Then argument above also holds for finite F with exactly the same construction and so if F is finite and such that $\mu * \nu(F) = 1$ then there must exist finite sets G and H such that $\mu(G) = 1$ and $\nu(H) = 1$.

2

2.1

(1) $\Leftrightarrow$ (2)
Suppose u and ν are σ-finite *positive* measures on $(\Omega, \mathcal{F})$. Then suppose that $\mu << \nu$ and $\nu << \mu$. Then for $A \in \mathcal{F}$ we have that $\mu(a) = 0 \Rightarrow \nu(A) = 0$ and $\nu(A) = 0 \Rightarrow \mu(A)$. That is to say that $\nu(A) = 0 \Leftrightarrow \mu(A) = 0$. That is to say that ν and μ have the same null sets. This argument is symmetric so it is clear that the reverse implication also holds.
(1) $\Longrightarrow$ (3)
We now wish to show that there is an $\mathcal{F}$-measurable function g that satisfies $0 < g(\omega) < +\infty$ at each $\omega \in \Omega$ and is such that $\nu(A) = \int_A g d\mu$ for all $A \in \mathcal{F}$.
The Radon-Nikodym theorem guarantees that there must exist a *unique* $g \in L^1(\mu)$ such that

$$\nu(A) = \int_A f d\mu. \tag{21}$$

[1] Note that the reason we could split up the integrals to a product in both cases is because $\int \chi_H(y)\nu(y)$ is a constant with respect to x.

However, f is unique only up to sets of measure 0. Due to the fact that $\nu(A) \geq 0$ for all $A \in \mathcal{F}$ we must have that $\int_A f d\mu \geq 0$ for all $A \in \mathcal{F}$ and hence by proposition 2.20 (4) $f \geq 0$ on every $A \in \mathcal{F}$ μ-a.e.
That is to say that $f < 0$ on a set of measure 0. So we write

$$h(\omega) = \begin{cases} 1 & \omega \in [-\infty, 0) \\ f & \text{otherwise.} \end{cases} \tag{22}$$

Notice that $[-\infty, 0) \in \mathcal{F}$ and so h is a $\mathcal{F}$ measurable function.
Now

$$\left(\frac{d\mu}{d\nu}\right)\left(\frac{d\nu}{d\mu}\right) = 1 \tag{23}$$

almost everywhere implies that

$$\left(\frac{d\mu}{d\nu}\right) = 0 \text{ iff } \left(\frac{d\nu}{d\mu}\right) = \infty \tag{24}$$

and vica versa. However, we assumed that $\nu \sim \mu$ (1) and hence $f = 0$ and $f = \infty$ also only happen on a set of measure 0 so we can write

$$g(\omega) = \begin{cases} 1 & \omega \in [-\infty, 0] \cup \{\infty\} \\ f & \text{otherwise.} \end{cases} \tag{25}$$

and because $[-\infty, 0] \in \mathcal{F}$ and $\{\infty\} \in \mathcal{F}$ so we have that g is $\mathcal{F}$ measurable. Further $0 < g(\omega) < \infty$.
(3) $\Longrightarrow$ (1).
Now suppose that there is a $\mathcal{F}$ measurable function $0 < g(\omega) < \infty$ such that $\nu(A) = \int_A g d\mu$.
Suppose $\nu(A) = 0$ then

$$\begin{aligned} 0 = \nu(A) &= \int_A g d\mu \qquad (26) \\ &= \int \chi_A g d\mu \qquad (27) \end{aligned}$$

Now as $\chi_A g >= 0$ and μ is a positive measure then by theorem 2.18 (3) we must that $\chi_A g = 0$, $\mu-$a.e. As $g > 0$ we have that $\chi_A = 0$ μ-a.e. which means that $\mu(A) = 0$. So $\mu \ll \nu$.
We claim that $\nu \ll \mu$ as well. Suppose $\mu(A) = 0$ then note that as g is a measurable function we have that by Theorem 2.11 $g(\omega) = \lim_{n \longrightarrow \infty} s_n(\omega)$ for all ω and a sequence of increasing simple functions. That is

$$\int_A g d\mu = \int_A \lim_{n \longrightarrow \infty} s_n d\mu. \tag{28}$$

Further by the monotone convergence theorem we have that

$$\int_A \lim_{n \longrightarrow \infty} s_n d\mu = \lim_{n \longrightarrow \infty} \int_A s_n d\mu. \tag{29}$$

Now for some simple function s we have that if $\mu(A) = 0$ then

$$\int_A s d\mu = \int_A \sum_{k=1}^{N} \alpha_k \chi_{B_k} d\mu \tag{30}$$

$$= \sum_{k=1}^{N} \alpha_k \int_A \chi_{B_k} d\mu \tag{31}$$

$$= \sum_{k=1}^{N} \alpha_k \int \chi_{B_k \cap A} d\mu \tag{32}$$

$$\leq \sum_{k=1}^{N} \alpha_k \int \chi_A d\mu \tag{33}$$

$$= \sum_{k=1}^{N} \alpha_k 0 \tag{34}$$

$$= 0 \tag{35}$$

thus if $\mu(A) = 0$ then $\int_A s_n d\mu = 0$ for all n and so $\lim_{n \longrightarrow \infty} \int_A s_n d\mu = 0$ and hence $\nu(A) = 0$. This means that $\nu \ll \mu$ as well.

2.2

Let $\{B_n\}_{n \in \mathbb{N}}$ be a covering of Ω by disjoint sets with $0 < \mu(B_n) < \infty$ for all n. Such a covering exists because μ is σ-finite. Now select a sequence of constants $\{\alpha_n\}_{n \in \mathbb{N}}$ such that $\alpha_n > 0$ for all n and

$$\sum_{n=1}^{\infty} \alpha_n = 1. \tag{36}$$

For example we could select $\alpha_n = 2^{-n}$.
We claim that the function $\nu : \mathcal{F} \longrightarrow \mathbb{R}$ defined by

$$\nu(A) = \sum_{n=1}^{\infty} \alpha_n \frac{\mu(A \cap B_n)}{\mu(B_n)} \tag{37}$$

is a probability measure with the same null sets as μ. Firstly we show that it is a measure, that is we show σ-additivity.
For a collection of disjoint sets $\{A_n\}_{n \in \mathbb{N}}$ we have that

$$\nu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{k=1}^{\infty} \alpha_k \frac{\mu\left(\bigcup_{n=1}^{\infty} A_n \cap B_k\right)}{\mu(B_k)} \tag{38}$$

and by σ-additivity of μ we get

$$\sum_{k=1}^{\infty} \alpha_k \frac{\mu\left(\bigcup_{n=1}^{\infty} A_n \cap B_k\right)}{\mu(B_k)} = \sum_{k=1}^{\infty} \alpha_k \frac{\sum_{n=1}^{\infty} \mu\left(A_n \cap B_k\right)}{\mu(B_k)} \tag{39}$$

$$= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \alpha_k \frac{\mu\left(A_n \cap B_k\right)}{\mu(B_k)} \tag{40}$$

$$= \sum_{n=1}^{\infty} \nu(A_n). \tag{41}$$

The interchange of the order of summation can be justified by the fact that $\mu(A_n \cap B_k) \geq 0$ and

$$\sum_{k=1}^{\infty} \alpha_k \frac{\sum_{n=1}^{\infty} \mu\left(A_n \cap B_k\right)}{\mu(B_k)} < \infty \tag{42}$$

which we prove now (by proving ν is a probability measure).
See that

$$\nu(\Omega) = \sum_{k=1}^{\infty} \alpha_k \frac{\mu(\Omega \cap B_k)}{\mu(B_k)} \tag{43}$$

$$= \sum_{k=1}^{\infty} \alpha_k \frac{\mu(B_k)}{\mu(B_k)} \tag{44}$$

$$= \sum_{k=1}^{\infty} \alpha_k \tag{45}$$

$$= 1. \tag{46}$$

We now just have to show that μ and ν share the same null sets.
Suppose $\mu(A) = 0$ then

$$\mu(A) = \sum_{k=1}^{\infty} \alpha_k \frac{\mu(A \cap B_k)}{\mu(B_k)} \tag{47}$$

$$\leq \sum_{k=1}^{\infty} \alpha_k \frac{\mu(A)}{\mu(B_k)} \tag{48}$$

$$= 0 \tag{49}$$

and thus $\nu \ll \mu$. Now suppose $\nu(A) = 0$ then

$$\sum_{k=1}^{\infty} \alpha_k \frac{\mu(A \cap B_k)}{\mu(B_k)} = 0 \tag{50}$$

and as $\mu(B_k) < \infty$ for all k this implies that $\mu(A \cap B_k) = 0$ for all k. Now

$$0 = \sum_{k=1}^{\infty} \mu(A \cap B_k) \tag{51}$$

$$= \mu\left(\bigcup_{k=1}^{\infty} (A \cap B_k)\right) \tag{52}$$

$$= \mu\left(A \cap \bigcup_{k=1}^{\infty} B_k\right) \tag{53}$$

$$= \mu(A \cap \Omega) \tag{54}$$

$$= \mu(A). \tag{55}$$

Thus $\mu \ll \nu$.
So ν is a finite measure on $(\Omega, \mathcal{F})$ which is equivalent to μ.

3

3.1

Firstly note that the definition of the characteristic function is

$$\hat{\mu}_X(u) = \mathbb{E}[\exp(i\langle X, u\rangle)] \tag{56}$$

and so for the random vector cX with $c \in \mathbb{R}$ we have that

$$\begin{aligned}
\hat{\mu}_{cX}(u) &= \mathbb{E}[\exp(i\langle cX, u\rangle)] &(57)\\
&= \mathbb{E}[\exp(i\langle X, cu\rangle)] &(58)\\
&= \hat{\mu}_X(cu) &(59)
\end{aligned}$$

3.2

By definition we have that

$$\hat{\mu}(\mathbf{u}) = \int \exp(i\langle \mathbf{x}, \mathbf{u}\rangle) d\mathbb{P}_X(\mathbf{x}). \tag{60}$$

We can say that

$$\frac{d^\alpha \hat{\mu}(\mathbf{u})}{d\mathbf{u}^\alpha} = \frac{d^\alpha}{d\mathbf{u}^\alpha} \int \exp(i\langle \mathbf{x}, \mathbf{u}\rangle) d\mathbb{P}_X(\mathbf{x}). \tag{61}$$

We wish to justify taking this derivative through the integral sign. To do this we use an extension of corollary 2.28 (2) from the notes.

Claim
If the partial derivative

$$\frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) \tag{62}$$

exists for all $(\mathbf{x}, \mathbf{u}) \in X \times [a, b]^d$ and if there is a function $g \in L^1(\mu)$ such that

$$\left| \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) \right| \leq g \tag{63}$$

for every $\mathbf{x} \in X$ and $\mathbf{u} \in (a, b)^d$ then

$$\frac{d^\alpha}{d\mathbf{u}^\alpha} \int f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) = \int \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) \qquad \text{for } \mathbf{u} \in (a, b)^d \tag{64}$$

Proof
We induct on $|\alpha|$. For $|\alpha| = 0$ the result is degenerate. Suppose it is true for $|\alpha|$ we wish to show that it is true for $|\alpha| + 1$. If it is true for $|\alpha|$ then we have that

$$\frac{d^\alpha}{d\mathbf{u}^\alpha} \int f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) = \int \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) \qquad \text{for } \mathbf{u} \in (a, b)^d \tag{65}$$

then for $|\alpha| + 1$ with a derivative with respect to the ith component we would have

$$\frac{d}{du_i} \frac{d^\alpha}{d\mathbf{u}^\alpha} \int f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) = \frac{d}{du_i} \int \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) \qquad \text{for } \mathbf{u} \in (a, b)^d \tag{66}$$

but by corollary 2.28 (2) we have

$$\frac{d}{du_i} \int \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) = \int \frac{\partial}{\partial u_k} \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) \tag{67}$$

$$= \int \frac{\partial^{\alpha+1}}{\partial \mathbf{u}^{\alpha+1}} f(\mathbf{x}, \mathbf{u}) d\mu(\mathbf{x}) \qquad \text{for } \mathbf{u} \in (a, b)^d \tag{68}$$

and so we have the result by induction. [a]

[a] We abused notation with $|\alpha| + 1$ and $\alpha + 1$ but by $\alpha + 1$ we mean we increase the order of the derivative of the ith variable. Also note that we assume that we can interchange the orders of differentiation (or combine them). This is only allowed if f has continuous partial derivatives of the appropriate order. For this case we assume this is ok.

We have that

$$\frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} \exp(i\langle \mathbf{x}, \mathbf{u} \rangle) = i^{|\alpha|} \prod_{k=1}^{d} x_k^{\alpha_k} \exp(i\langle \mathbf{x}, \mathbf{u} \rangle) \tag{69}$$

and that

$$\left| \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} \exp(i\langle \mathbf{x}, \mathbf{u} \rangle) \right| \leq \underbrace{\prod_{k=1}^{d} |x_k|^{\alpha_k}}_{\circledast} . \tag{70}$$

Now because

$$\mathbf{E}\left(\prod_{k=1}^{d} |X_k|^{\alpha_k} \right) = \int \prod_{k=1}^{d} |x_k|^{\alpha_k} \, d\mathbb{P}_X(\mathbf{x}) < \infty \tag{71}$$

then $\circledast \in L^1(\mathbb{P}_X)$ and so we can apply our claim (the DCT) to get

$$\frac{\partial^\alpha \hat{\mu}(\mathbf{u})}{\partial \mathbf{u}^\alpha} = \int \frac{\partial^\alpha}{\partial \mathbf{u}^\alpha} \exp(i\langle \mathbf{x}, \mathbf{u}\rangle) d\mathbb{P}_X(\mathbf{x}) \tag{72}$$

$$= \int i^{|\alpha|} \prod_{k=1}^{d} x_k^{\alpha_k} \exp(i\langle \mathbf{x}, \mathbf{u}\rangle) d\mathbb{P}_X(\mathbf{x}) \tag{73}$$

$$= i^{|\alpha|} \int \prod_{k=1}^{d} x_k^{\alpha_k} \exp(i\langle \mathbf{x}, \mathbf{u}\rangle) d\mathbb{P}_X(\mathbf{x}) \tag{74}$$

and so

$$\frac{\partial^\alpha \hat{\mu}(\mathbf{u})}{\partial \mathbf{u}^\alpha}\Big|_{\mathbf{u}=\mathbf{0}} = i^{|\alpha|} \int \prod_{k=1}^{d} x_k^{\alpha_k} \exp(i\langle \mathbf{x}, \mathbf{0}\rangle) d\mathbb{P}_X(\mathbf{x}) \tag{75}$$

$$= i^{|\alpha|} \int \prod_{k=1}^{d} x_k^{\alpha_k} d\mathbb{P}_X(\mathbf{x}) \tag{76}$$

$$= i^{|\alpha|} \mathbb{E}(X^\alpha) \tag{77}$$

3.3

Notice that as the distribution is symmetric we have that

$$\hat{\mu}(u) = \int e^{iux} f(x) dx \tag{78}$$

$$= \int f(x)(\cos(ux) + i\sin(ux)) dx \tag{79}$$

$$= \int f(x) cos(ux) dx \tag{80}$$

Now notice that

$$\left| \frac{\partial}{\partial u} f(x)\cos(ux) \right| = |xf(x) sin(xu)| \tag{81}$$

Now notice that

$$\int |xf(x)\sin(ux)| dx = \underbrace{\int_{-K}^{K} |xf(x)\sin(ux)| dx}_{\leq M_1} + \int_{-\infty}^{-K} |xf(x)\sin(ux)| dx + \int_{K}^{\infty} |xf(x)\sin(ux)| dx \tag{82}$$

for some constant M_1 and a constants K and J chosen such that

$$\int_{K}^{\infty} |xf(x)\sin(ux)| dx \leq \int_{K}^{\infty} \frac{|\sin(ux)|}{Jx} dx. \tag{83}$$

We can therefore say that

$$' \int |xf(x)\sin(ux)| dx \leq M_1 + 2\int_{K}^{\infty} \frac{|\sin(ux)|}{Jx} dx. \tag{84}$$

Now appealing to a result in special functions we can say that

$$\int_{K}^{\infty} \frac{|\sin(ux)|}{Jx} dx = \lim_{x \longrightarrow \infty} \mathrm{Si}(ux)\,\mathrm{sgn}(\sin(ux)) - \mathrm{Si}(uK)\,\mathrm{sgn}(\sin(uK)) \tag{85}$$

and it turns out that $\frac{1}{J}\lim_{x\longrightarrow\infty}\mathrm{Si}(ux)\,\mathrm{sgn}(\sin(ux))<\infty$ so long as $u\neq 0$ which means that

$$\left|\frac{\partial}{\partial u}f(x)\cos(ux)\right|\in L^1(\mu) \tag{86}$$

for every $u\neq 0$. We can then apply the dominated convergence theorem corollary 2.28 (2) to say that $\hat{\mu}(u)$ is differentiable for all $u\neq 0$ and by extension, sufficiently close to 0.

4

Let μ be the binomial distribution with n trials and probability of success p, that is $\mu=\mathrm{Bin}(n,p)$, and let ν be the Poisson distribution with mean $\lambda>0$.

4.1

We wish to verify that $\hat{\mu}(u)=(1-p+pe^{iu})^n$. Because the binomial distribution is just the convolution of identical independent Bernoulli distributions then we just have to verify that $(1-p+pe^{iu}$ is the characteristic function for $Bernoulli(p)$.
If ν is the Bernoulli measure and X has law ν then

$$\begin{aligned}\hat{\nu}(u)&=\mathbb{E}[\exp(iuX)] &(87)\\ &=\sum_{k\in\{0,1\}}e^{iuk}\nu_X(k) &(88)\\ &=pe^{iu}+(1-p). &(89)\end{aligned}$$

Then by repeated application of the convolution theorem we get that $\hat{\mu}(u)=(1-p+pe^{iu})^n$.

4.2

We wish to verify that $\hat{\nu}(u)=\exp(\lambda(e^{iu}-1))$. The probability mass function of the Poisson distribution is

$$\frac{\lambda^k}{k!}e^{-\lambda} \tag{90}$$

and thus

$$\begin{aligned}\mathbb{E}[\exp(iuX)]&=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}e^{iuk} &(91)\\ &=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}(e^{iu})^k &(92)\\ &=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^{iu})^k}{k!} &(93)\\ &=e^{-\lambda}e^{\lambda e^{iu}} &(94)\\ &=e^{\lambda(e^{iu}-1)} &(95)\end{aligned}$$

4.3

We wish to show that if p_n is a sequence in $[0,1]$ such that $p_n \downarrow 0$ and $np_n \longrightarrow \lambda$ then $\mu_n \longrightarrow \nu$ in the weak sense where $\mu_n = \text{Bin}(n, p_n)$. Let $f \in C_b$ then

$$\lim_{n \longrightarrow \infty} \sum_{k=0}^{n} f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k} = \lim_{n \longrightarrow \infty} \sum_{k=0}^{\infty} \chi_{k \leq n} \cdot f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k} \tag{96}$$

$$= \sum_{k=0}^{\infty} \lim_{n \longrightarrow \infty} \chi_{k \leq n} \cdot f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k}. \tag{97}$$

The interchange of the order of the limit and the sum is justified by the uniform convergence of the sum. To see this let $M = \sup_{k \in \mathbb{N}^0} f(k)$ (which exists because $f \in C_b$) and then note that

$$\sum_{k=0}^{\infty} \chi_{k \leq n} \cdot f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k} \leq \sum_{k=0}^{\infty} \chi_{k \leq n} \cdot M \binom{n}{k} p_n^k (1-p_n)^{n-k} \tag{98}$$

$$= M < \infty \tag{99}$$

and so by the Weierstrass M test the series converges uniformly.
Now as $np_n \longrightarrow \lambda$ or $p_n \longrightarrow \frac{\lambda}{n}$ we get

$$\sum_{k=0}^{\infty} \lim_{n \longrightarrow \infty} \chi_{k \leq n} \cdot f(k) \binom{n}{k} p_n^k (1-p_n)^{n-k} = \sum_{k=0}^{\infty} \lim_{n \longrightarrow \infty} \chi_{k \leq n} \cdot f(k) \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^n (1-p_n)^{-k} \tag{100}$$

$$= \sum_{k=0}^{\infty} \lim_{n \longrightarrow \infty} \chi_{k \leq n} \cdot f(k) \frac{n^k + O(n^{k-1})}{k!} p_n^k (1-p_n)^n (1-p_n)^{-k} \tag{101}$$

$$= \sum_{k=0}^{\infty} f(k) \lim_{n \longrightarrow \infty} \frac{n^k + O(n^{k-1})}{k!} p_n^k (1-p_n)^n \underbrace{(1-p_n)^{-k}}_{\longrightarrow 0} \tag{102}$$

$$= \sum_{k=0}^{\infty} f(k) \lim_{n \longrightarrow \infty} \underbrace{\frac{n^k + O(n^{k-1})}{k!} p_n^k}_{\longrightarrow \frac{\lambda^k}{k!}} (1-p_n)^n \tag{103}$$

$$= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} f(k) \lim_{n \longrightarrow \infty} \underbrace{(1 - \frac{\lambda}{n})^n}_{\longrightarrow e^{-\lambda}} \tag{104}$$

$$= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} f(k). \tag{105}$$

This proves the weak convergence.

4.4

This argument holds whether one takes the integral (sum) or not. So $\mu_n(\{k\}) \longrightarrow \nu(\{k\})$ for all $k \in \mathbb{N}^0$.

5

5.1

We wish to show that $\mathbb{P}(B_n) = 1/2$ for every $n \geq 1$. Note that

$$B_n = \bigcup_{k=0}^{2^{n-1}-1} \left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right) \tag{106}$$

and thus

$$\begin{aligned}
\mathbb{P}(B_n) &= \mathbb{P}\left(\bigcup_{k=0}^{2^{n-1}-1} \left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right)\right) && (107)\\
&= \sum_{k=0}^{2^{n-1}-1} \mathbb{P}\left(\left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right)\right) && (108)\\
&= \sum_{k=0}^{2^{n-1}-1} \frac{1}{2^n} && (109)\\
&= 2^{n-1}\frac{1}{2^n} && (110)\\
&= \frac{1}{2}. && (111)
\end{aligned}$$

5.2

We now wish to show that the sequence of events B_n form an infinite sequence of independent events. Take a finite subset $J \subset \mathbb{N}$ with $|J| = m$ and $\max J = r$ then

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{n\in J} B_n\right) &= \mathbb{P}\left(\bigcap_{n\in J} \bigcup_{k=0}^{2^{n-1}-1} \left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right)\right) && (112)\\
&= \mathbb{P}\left(\bigcup_{k=0}^{2^{r-m}-1} \left[\frac{2k}{2^r}, \frac{2k+1}{2^r}\right)\right) && (113)\\
&= \sum_{k=0}^{2^{r-m}-1} \mathbb{P}\left(\left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right)\right) && (114)\\
&= 2^{r-m}\frac{1}{2^r} && (115)\\
&= \frac{1}{2^m} && (116)\\
&= \prod_{n\in J} \mathbb{P}\left(B_n\right) && (117)
\end{aligned}$$

and so the sequence of events B_n form an infinite sequence of independent events.

5.3

We wish to show / argue that the probability that a randomly sampled number ω will have the sequence 5825 occur infinitely often in its decimal expansion is 1.

We use the Borel-Cantelli lemma. Ignoring possible overlaps (on the 5s) we can see that we can break any decimal expansion of ω up into blocks of 4 digits.
Then by we can define E_i as the probability of obtaining 5285 in the i-th block possition. By the same argument as above these events are independent.
The for any i we have that $\mathbb{P}(E_i) = \frac{1}{10000}$ (the same argument as above applied to a decimal expansion). Then clearly

$$\mathbb{P}(E_i) = \infty. \tag{118}$$

By the Borel-Cantelli lemma this implies

$$\mathbb{P}(\limsup_n(E_n) = 1. \tag{119}$$

Now

$$\limsup_n(E_n) = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} E_j \tag{120}$$

can be intuitively read as E_j happens infinitely often. Which is to say that 5285 occurs *blockwise* in the expansion of ω infinitely often. Clearly as allowing for overlaps allows for more configurations then the probability is 1 (it can be no more).

6

We wish to prove that if A and B are i.i.d random variables with finite variance and $A+B$ and $A-B$ are independent then A and B are normal. This result is actually called Bernstein's Theorem. (Actually there are lots of things called Bernstein's Theorem, but this is one of them.)[2].
Firstly write $X = (A - \mathbb{E}(A))/\sqrt{\mathbb{E}(A^2)}$ and $Y = (B - \mathbb{E}(B))/\sqrt{\mathbb{E}(B^2)}$. $\mathbb{E}(A)$, $\mathbb{E}(A^2)$, $\mathbb{E}(B)$ and $\mathbb{E}(B^2)$ exist because $\mathrm{Var}(A) < \infty$ and $\mathrm{Var}(B) < \infty$.
Notice that $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = 1$.
Notice that we can write

$$X = \frac{1}{2}((X+Y) + (X-Y)). \tag{121}$$

So if we write $\hat{\mu_X}(u)$, $\hat{\mu_Y}(u)$,$\mu_{X\hat{+}Y}(u)$ and $\mu_{X\hat{-}Y}(u)$ as the characteristic functions of X, Y, $X+Y$ and $X-Y$ respectively we can write.

$$\hat{\mu_X}(u) = \mu_{X\hat{+}Y}(\frac{u}{2})\mu_{X\hat{-}Y}(\frac{u}{2}) \tag{122}$$

because $X+Y$ and $X-Y$ are independent. Now because X and Y are independent we can further write

$$\hat{\mu_X}(u) = \hat{\mu_X}(\frac{u}{2})^2\hat{\mu_Y}(u/2)\hat{\mu_Y}(-\frac{u}{2}). \tag{123}$$

Now as X and Y are identically distributed they have the same characteristic function and so we can say $\hat{\mu_X} \equiv \hat{\mu_Y}$ and hence from now on we just write $\hat{\mu}$.

$$\hat{\mu}(u) = \hat{\mu}(\frac{u}{2})^3\hat{\mu}(-\frac{u}{2}). \tag{124}$$

[2]This proof is based off ideas from
`http://math.stackexchange.com/questions/556030/x-and-y-i-i-d-xy-and-x-y-independent-mathbbex-0-and-mathbb`

Lets define $\Psi(u) = \log(\hat{\mu}(u))$. In this case the functional equation reduces to

$$\Psi(u) = 3\Psi(\frac{u}{2}) + \Psi(-\frac{u}{2}) \tag{125}$$

and

$$\Psi(-u) = 3\Psi(-\frac{u}{2}) + \Psi(\frac{u}{2}) \tag{126}$$

Now by defining $\Xi(u) = \Psi(u) - \Psi(-u)$. We can reduce the functional equation to $\Xi(u) = 2\Xi(\frac{u}{2})$. Then

$$\Xi(\frac{u}{2^n}) = 2^n \Xi(u) \tag{127}$$

and by rearranging and dividing by t we get

$$\frac{\Xi(u)}{u} = \frac{\Xi(\frac{u}{2^n})}{\frac{u}{2^n}}. \tag{128}$$

Now if $\mathbb{E}[X] = \alpha$ and $\mathbb{E}[X^2] = \beta$ (these values exist because of the assumption of finite variance) we get that $\hat{\mu}'(0) = 0$ and $\hat{\mu}''(0) = -1$ by the results of 3.2 of this assignment. Now as $\hat{\mu}$ is twice differentiable then so are Ξ and Ψ. Now $\Xi'(u) = 2\Psi'(u)$ and $\Psi'(u) = \frac{\hat{\mu}'(u)}{\hat{\mu}(u)}$. Thus $\Xi'(0) = 0$ and since

$$\frac{\Xi(u)}{u} = \frac{\Xi(\frac{u}{2^n})}{\frac{u}{2^n}} \tag{129}$$

and

$$\lim_{n \longrightarrow \infty} \frac{\Xi(\frac{u}{2^n})}{\frac{u}{2^n}} = \Xi'(0) \tag{130}$$

and thus $\Xi[u] = i\alpha u$.
Now

$$\Psi''(u) = \frac{\hat{\mu}''(u)\hat{\mu}(u) - (\hat{\mu}'(u))^2}{\hat{\mu}(u)^2} \tag{131}$$

and at zero this means $\Psi''(0) = -1$. By using the previous result for $\Xi'[u] = 0$ along with its definition we have that

$$\Psi(u) = 4\Psi(\frac{u}{2}) \tag{132}$$

and so

$$\frac{\Psi(u)}{u^2} = \frac{4}{\frac{u^2}{4^n}} \Psi(\frac{u}{2^n}). \tag{133}$$

Now

$$\lim_{n \longrightarrow \infty} \frac{4}{\frac{u^2}{4^n}} \Psi(\frac{u}{2^n}) = \frac{1}{2}\Psi''(0) \tag{134}$$

through one application of L'Hopitals rule.
Thus $\Psi(u) = -\frac{1}{2}(u^2)$ which means that $\hat{\mu}(u) = \exp(-\frac{1}{2}(u^2))$. Notice that this is the characteristic function of a standard normal distribution and so X and Y both have standard normal distributions. By rescaling we get that A and B are also normally distributed (just with different parameters).
It is not hard to see that this result generalises to larger collections of random variables. For a proof of that result see *Lukacs, Eugene; King, Edgar P. (1954). "A Property of Normal Distribution" The Annals of Mathematical Statistics 25 (2): 389 —394.*