MEASURE THEORETICAL PROBABILITY I HOMEWORK 3

ANDREAS PETROU-ZENIOU; CHENG PENG; TOMMENIX YU; CONROY LAU
STAT 38100
DUE FEB 1

Discussed with classmates.

Exercise 0.

Proof.

Riemann integrable $\Rightarrow$ Lebesgue integrable:

Since f is Riemann integrable, so is f^+ and f^-. Thus we can WLOG assume f is positive and prove that the Lebesgue integral is finite.

f is Riemann integrable means that for any $\varepsilon > 0$, there exists partition

$$P = P(\varepsilon) = \{p_0, \dots, p_n\}$$

where $p_0 = a$, $p_n = b$ such that

$$\sum_{i=0}^{n-1} \left(\max_{x\in[p_i,p_{i+1}]} f(x) - \min_{x\in[p_i,p_{i+1}]} f(x) \right) (p_{i+1} - p_i) < \varepsilon.$$

Now since f is bounded

But this means that we can find a corresponding non-negative simple function

$$s_P(x) = \min_{x\in[p_i,p_{i+1}]} f(x) \quad \text{for} \quad x \in [p_i, p_{i+1}]$$

such that

$$\int_a^b f dx \geq \int_a^b s_P dx$$

and

$$\int_a^b f dx - \int_a^b s_P dx \leq \varepsilon$$

which since ε is arbitrary we can take the limit of $\varepsilon \to 0$ (note that s_P depends on it also) and get that

$$\int_b^a f d\lambda = \sup_{u\in SF(f)^+} \int_a^b u d\lambda \geq \sup_{P(\varepsilon)} \int_a^b s_P d\lambda = \int_b^a f dx \quad \text{(Riemann)}$$

And since for all t, s simple with $t \geq f \geq s$, we know that

$$\int_a^b s d\lambda \leq \int_a^b f d\lambda \leq \int_a^b t d\lambda$$

thus taking sup on s we get

$$\int_a^b \sup_s s d\lambda \leq \int_a^b t d\lambda$$

for any t. Thus by a similar construction of s_P we can construct t_P such that

$$\inf_{t_P} \int_a^b t_P d\lambda = \int_b^a f dx$$

and hence

$$\int_b^a f dx \leq \sup_{u \in SF(f)^+} \int_a^b u d\lambda \leq \int_b^a f dx$$

so

$$\int_a^b f d\lambda = \int_b^a f dx$$

which is finite. Thus f is integrable (by argument on top). $\square$

Exercise 1. *Prob 1.*

Proof.

Use counter example to show these conditions are necessary.

Non-negativity condition in the Monotone Convergence Theorem:

Let $f : \mathbb{R} \to \mathbb{R}$ be such a function such that

$$f(x) = \frac{1}{1+x^2}$$

which is a well integrable function. Now define

$$f_n(x) = \begin{cases} f(x) & x \le n \\ -1 & x > n \end{cases}$$

then f_n is an increasing sequence of function which converges point wise to f. But then by computation

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n dx = \lim_{n\to\infty} -\infty = -\infty \neq \pi = \int_{\mathbb{R}} f dx.$$

In fact, this is a counter-example to the first 3 conditions.

Non-negativity condition in Fatou's Lemma:

Since the above function has a limit, it is equal to the liminf. Then

$$\liminf_n \int_{\mathbb{R}} f_n dx = \liminf_n 0 = 0 \geq \pi = \int_{\mathbb{R}} f dx = \int_{\mathbb{R}} \liminf_n f_n dx$$

which contradicts to the claim for Fatou's lemma.

Dominated condition to Dominated convergence theorem:

Note that $|f_n|$ is not dominated, and we know from above that we cannot pass limit.

Non-negativity/integrable condition for Fubini's

We can just use a countable measure space $(\mathbb{N}^+ \times \{1,2\}, \mathcal{F}, \mu)$ where $\mathcal{F}$ is just the power set of $\mathbb{N}$ and μ assigns weight uniformly, and we can WLOG assume $\mu(\{k\}) = 1$. Also, the product space the self product of the above space.

Now we choose $a_{i,n} = \dfrac{(-1)^i}{n}$, $n-1, 2, ...,\ i \in \{1,2\}$, which is a measurable function since every function is measurable for the power-set σ-algebra.

However, first integrating in i and then in n yields

$$\sum_{i=1}^{2} \sum_{n=1}^{\infty} a_{i,n} = \infty - \infty$$

which does not converge.

But do it in the other direction we get

$$\sum_{n=1}^{\infty}\sum_{i=1}^{2} a_{i,n} = 0$$

A even more explicit counter example is the following. Here nothing is undefined.

Source: https://math.jhu.edu/~jmb/note/nofub.pdf

The example from source shows a multivariate function (only 2 directions) can be defined and all the expressions needed in Fubini is well defined, and even finite, but the ordered integration just don't coincide in value. Even the source cannot explain this too well. I think that it's because the function (if cut off 0) is the standard counter example for Gateaux differentiable not being Frechet differentiable, so really at around 0, the order of direction in which you look at it matters. This is also reflected in the integration, as we'll see now.

Let $X = [0,2], Y = [0,1]$, the σ-algebra being the Borel one and the usual Lebesgue measure. Define

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{(x^2+y^2)^3} & (x,y) \neq (0,0) \\ 0 & (x.y) = (0,0) \end{cases}.$$

Then we compute. We first use the change of variable $u = x^2 + y^2$ and get

$$\int_0^1 f(x,y)dy = \int_{x^2}^{x^2+1} \frac{x(2x^2-u)}{2u^3}du = \int_{x^2}^{x^2+1}\left(\frac{x^3}{u^3} - \frac{x}{2u^2}\right)du = \frac{x}{2(x^2+1)^2}$$

and hence

$$\int_0^2\int_0^1 f(x,y)dydx = \int_0^2 \frac{x}{2(x^2+1)^2}dx = \frac{1}{5}.$$

Yet for the other direction we have

$$\int_0^2 f(x,y)dx = \int_{y^2}^{y^2+4} \frac{y(u-2y^2)}{2u^3}du = -\frac{2y}{(4+y^2)^2}$$

and

$$\int_0^1\int_0^2 f(x,y)dxdy = \int_0^1 -\frac{2y}{(4+y^2)^2}dy = -\frac{1}{20}.$$

Thus

$$\int_0^2\int_0^1 f(x,y)dydx \neq \int_0^1\int_0^2 f(x,y)dxdy$$

and Fubini doesn't hold.

□

Exercise 2. *Prob 2.*

Proof.

the derivative is integrable:

Since $g_t(x) := \frac{d}{dt} f(x,t)$ (well defined) has absolutely value less than h, thus so is g^+ and g^-, which means g is integrable if it is measurable. But it is measurable since by definition

$$g_t(x) = \lim_{n\to\infty} \frac{f(x, t+1/n) - f(x,t)}{1/n}$$

which is a limit of measurable functions. Thus, g is integrable.

$\int_X f(t,x)d\mu$ is differentiable with $\frac{d}{dt}\int_X f(t,x)d\mu = \int_X \frac{d}{dt} f(t,x)d\mu$:

We show that it is differentiable by showing that the derivative quotient is well-defined, that the limit exists. This, as we'll see, gives us exactly the value of it's derivative.

The derivative quotient of $\int_X f(t,x)d\mu$ is (formally written as):

$$\frac{d}{dt}\int_X f(t,x)d\mu = \lim_{s\to 0} \frac{\int_X f(t+s,x)d\mu - \int_X f(t,x)d\mu}{s} = \lim_{s\to 0}\int_X \frac{f(t+s,x) - f(t,x)}{s} d\mu$$

by linearity. So we only need to define

$$h_s(t,x) := \frac{f(t+s,x) - f(t,x)}{s}$$

for $s > 0$, and prove that we can pass the limit when $s \to 0$. To do this we note that it is measurable because $f(t,x)$ is integrable, hence measurable, and $h_s(t,x)$ is a sum of measurable functions, times a constant, which is still measurable.

Now we show that there exists integrable function k such that $|h_s| \le k$. To do this we just show for all s sufficiently small, it does not matter since the limit we're taking is to 0.

$$k(t,x) = h(x) + |f(t,x)|$$

which is integrable because both are non-negative integrable functions. Thus, since h_s are measurable and is dominated by k, we get that

$$\lim_{s\to 0}\int_X h_s(t,x)d\mu = \int_X \lim_{s\to 0} h_s(t,x)d\mu$$

where we implicitly do the change of variable $s \to 1/n$ to fit the theorem. But observe that the above equality is nothing but

$$\frac{d}{dt}\int_X f(t,x)d\mu = \int_X \frac{d}{dt} f(t,x)d\mu$$

and we are done.

$\square$

Exercise 3. *Prob 3.*

Proof.

Say $Y = F(X)$. Consider now $P(F(X(\omega)) \leq y)$. We also know that F is continuous. Now, consider $\alpha = \sup\{F^{-1}([0, y])\}$. That is, 1 α is the supremum of all x such that $F(x) \leq y$. Because $[0, y]$ is closed, and F is continuous, this supremum is attained. That is, $F(\alpha) \leq y$, and for any $x \leq \alpha$, $F(x) \leq y$ while for any $x > \alpha$, $F(x) > y$. By the IVT, $F(\alpha) = y$.

Hence, $F(X(\omega)) \leq y$ if and only if $X(\omega) \leq \alpha$. Hence, $P(F(X(\omega)) \leq y) = P(X(\omega) \leq \alpha)$ which in turn equals y.

$\square$

Exercise 4. *Prob 4.*

Proof.

Source:https://en.wikipedia.org/wiki/Derangement

We say that a permutation is a derangement if the number of fixed point is 0, for example, $\{1,2\} \to \{2,1\}$ is a derangement. Actually it's the only derangement when $n = 2$. We denote the number of derangements in S_n by $!n$.

Now we derive a formula for the number of derangement in S_n. We use inclusion-exclusion to do this. Let $T_i \subset S_n$ be defined as

$$T_i : \left\{ S \in S_n \middle| S(i) = i \right\}$$

or that it's all the permutations that let i be a fixed point. Thus

$$!n = \#S_n \backslash \left(T_1 \cup \cdots \cup T_n\right)$$

where by inclusion-exclusion we have: ($|\cdot|$ is the cardinality)

$$\begin{aligned}
\left|T_1 \cup \cdots \cup T_n\right| &= \sum_{i=1}^{n} |T_i| - \sum_{i<j} |T_i \cup T_j| + \sum_{i<j<k} |T_i \cup T_j \cup T_k| - \dots \\
&= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - \dots \\
&= \frac{n!}{1!} - \frac{n!}{2!} + \frac{n!}{3!} - \dots \\
&= n! \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}
\end{aligned}$$

Hence

$$!n = |S_n| - \left|T_1 \cup \cdots \cup T_n\right| = n! - n! \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!} = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Now back to our random variable F_n, the number of fixed points of a rotation. We claim that there are

$$\binom{n}{k} \cdot !(n-k)$$

many permutations S that satisfies $F_n(S) = k$ because we first choose the k fixed points, then count all derangement of the other $n-k$ numbers.

Thus, since $|S_n| = n!$ we get that

$$\mathbb{P}(F_n = k) = \frac{\binom{n}{k} \cdot !(n-k)}{n!} = \frac{1}{n!} \frac{n!}{k! \cdot (n-k)!}(n-k)! \cdot \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} = \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!}$$

and we can compute that

$$\begin{aligned}
\mathbb{E}[F_n] &= \sum_{k=0}^{n} k\mathbb{P}(F_n = k) = \sum_{k=0}^{n} k\frac{1}{k!}\sum_{i=0}^{n-k}\frac{(-1)^i}{i!} = \sum_{k=1}^{n}\frac{1}{(k-1)!}\sum_{i=0}^{n-k}\frac{(-1)^i}{i!} \\
\text{(rearranging)} \ &= \sum_{i+k\le n, i\ge 0, k\ge 1}\frac{(-1)^i}{i!}\frac{1}{(k-1)!} = \sum_{i+k\le n, i\ge 0, k\ge 1}\frac{1}{(i+k-1)!}\binom{i+k-1}{i}(-1)^i \\
&= \sum_{m=1}^{n}\sum_{i+k=m, i\ge 0, k\ge 1}\frac{1}{(i+k-1)!}\binom{i+k-1}{i}(-1)^i \\
&= \sum_{m=1}^{n}\frac{1}{(m-1)!}\sum_{i=0}^{m-1}\binom{m-1}{i}(-1)^i = \sum_{m=1}^{n}\frac{1}{(m-1)!}\sum_{i=0}^{m-1}\binom{m-1}{i}(-1)^i 1^{m-1-i} \\
&= \sum_{m=1}^{n}\frac{1}{(m-1)!}(-1+1)^{m-1} = \sum_{m=1}^{n}\frac{1}{(m-1)!}0^{m-1} = 1.
\end{aligned}$$

Hence $\mathbb{E}[F_n] = 1$.

And it's clear that

$$\lim_{n\to\infty}\mathbb{P}(F_n = k) = \lim_{n\to\infty}\frac{1}{k!}\sum_{i=0}^{n-k}\frac{(-1)^i}{i!} = \frac{e^{-1}}{k!}.$$

$\square$

Exercise 5. *Prob 5.*

Proof.

$\mathbb{E}[X]$ exists since X is a no-negative measurable function. The right hand side exists because for each t, $\{X \geq t\}$ is measurable, and hence $\mathbb{P}(X \geq t)$ is non-negative. But since it's monotone, use result in pracice midterm we get that it is measurable, hence integrable.

(From the equality to the inequality below is direct, but there's a cool proof of the inequality.)

First we use the inequality (where the brackets are the closest integer that's smaller/larger than x)

$$\lfloor x \rfloor \leq x \leq \lceil x \rceil$$

and get

$$\mathbb{E}[\lfloor x \rfloor] \leq \mathbb{E}[x] \leq \mathbb{E}[\lceil x \rceil]$$

where we can compute

$$\mathbb{E}[\lfloor x \rfloor] = \int_\Omega \sum_{n=1}^{\infty} \mathbb{1}_{x \geq n} d\mathbb{P} = \sum_{n=1}^{\infty} \int_\Omega \mathbb{1}_{x \geq n} d\mathbb{P} = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n)$$

where we can exchange limit due to the linearity of integration. Similarly we have

$$\mathbb{E}[\lceil x \rceil] = \sum_{n=0}^{\infty} \mathbb{P}(X \geq n)$$

so we get

$$\sum_{n=1}^{\infty} \mathbb{P}(X \geq n) \leq \mathbb{E}[X] \leq \sum_{n=0}^{\infty} \mathbb{P}(X \geq n).$$

Now we prove the equality. Since

$$\mathbb{P}(X \geq t) = \int_\Omega \mathbb{1}_{X \geq t} d\mathbb{P}$$

which means

$$\int_0^\infty \mathbb{P}(X \geq t) dt = \int_0^\infty \int_\Omega \mathbb{1}_{X \geq t} d\mathbb{P} dt \overset{Fubini}{=} \int_\Omega \int_0^\infty \mathbb{1}_{X \geq t} dt d\mathbb{P} = \int_\Omega X(\omega) dt d\mathbb{P} = \mathbb{E}[X]$$

where we justify the use of Fubini by the non-negativity of $\mathbb{1}_{X \geq t}$ and that the spaces are σ-finite.

$\square$

Exercise 6. *Prob 6*

Proof.

Gaussian:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \sigma > 0.$$

$$\begin{aligned}
\mathbb{E}[f] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} d(x-\mu) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2}\int_{-\infty}^{\infty} e^{-\frac{y}{2\sigma^2}} dy + \int_{-\infty}^{\infty} \mu e^{-\frac{z^2}{2\sigma^2}} dz\right) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \left(\frac{1}{2}(-2\sigma^2)\int_{-\infty}^{\infty} e^{t} dt + \sqrt{2}\sigma\mu \int_{-\infty}^{\infty} e^{s^2} ds\right) \\
&= \frac{1}{\sigma\sqrt{2\pi}}(0 + \sqrt{2\pi}\sigma\mu) = \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}[f] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx - \mu^2 \\
&= \frac{\sigma\sqrt{2}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma\sqrt{2}t + \mu)^2 e^{-t^2} dt - \mu^2 \\
&= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 e^{-t^2} dt + 0 + \mu^2\sqrt{\pi}\right) - \mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt = \sigma^2
\end{aligned}$$

where the last step is just to change into dt^2 and do integration by parts.

Exponential:

$$\mathbb{E}[f] = \int_0^{\infty} x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}\int_0^{\infty} y e^{-y} dy = \frac{1}{\lambda}\left[-e^{-y} - ye^{-y}\right]_0^{\infty} = \frac{1}{\lambda}$$

$$\mathrm{Var}[f] = \int_0^\infty x^2\lambda e^{-\lambda x}dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}\int_0^\infty y^2e^{-y}dy - \frac{1}{\lambda^2}$$
$$= \frac{1}{\lambda^2}\left[-2e^{-y} - 2ye^{-y} - y^2e^{-y}\right]_0^\infty - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

<u>Gamma:</u>

$$\mathbb{E}[f] = \int_0^\infty x\frac{\lambda^r x^{r-1}}{\Gamma(r)}e^{-\lambda x}dx = \frac{\lambda^r}{\Gamma(r)}\int_0^\infty x^r e^{-\lambda x}dx$$
$$= \frac{\lambda^r}{\Gamma(r)}\int_0^\infty \left(\frac{t}{\lambda}\right)^r e^{-t}d\frac{t}{\lambda} = \frac{1}{\lambda\Gamma(r)}\int_0^\infty t^r e^{-t}dt = \frac{\Gamma(r+1)}{\lambda\Gamma(r)} = \frac{r}{\lambda}$$

$$\mathrm{Var}[f] = \frac{\lambda^r}{\Gamma(r)}\int_0^\infty x^{r+1}e^{-\lambda x}dx - \left(\frac{r}{\lambda}\right)^2 \overset{similarly}{=} \frac{\Gamma(r+2)}{\lambda^2\Gamma(r)} - \left(\frac{r}{\lambda}\right)^2 = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}$$

<u>uniform:</u>

$$\mathbb{E}[f] = \int_a^b x\frac{1}{b-a}dx = \frac{b+a}{2}$$
$$\mathrm{Var}[f] = \int_a^b x^2\frac{1}{b-a}dx - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

<u>Binomial:</u>

First, write $q := 1-p$ for convenience and

$$\mathbb{E}[f] = \sum_{k=0}^n k\binom{n}{k}p^kq^{n-k} = \sum_{k=1}^n k\binom{n}{k}p^kq^{n-k}$$
$$= \sum_{k=1}^n n\binom{n-1}{k-1}p^kq^{n-k} = np\sum_{k=1}^n \binom{n-1}{k-1}p^{k-1}q^{n-1-(k-1)} = np(p+q)^{n-1} = np$$

$$\mathrm{Var}[f] = \sum_{k=0}^n k^2\binom{n}{k}p^kq^{n-k} - n^2p^2 \overset{similarly}{=} np\sum_{k=1}^n k\binom{n-1}{k-1}p^{k-1}q^{n-1-(k-1)} - n^2p^2$$
$$= np\sum_{k=1}^n (k-1)\binom{n-1}{k-1}p^{k-1}q^{n-1-(k-1)} + np - n^2p^2$$
$$= np\sum_{k=1}^n (n-1)p\binom{n-2}{k-2}p^{k-2}q^{n-2-(k-2)} + np - n^2p^2 = np(n-1)p + np - n^2p^2$$
$$= np(1-p)$$

<u>X follows geometric distribution:</u>

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} k(1-p)^k p \\
&= p(1-p)\sum_{k=0}^{\infty}(1-p)^{k-1}k \\
&= p(1-p)\frac{d}{dp}(-\sum_{k=0}^{\infty}(1-p)^k) \\
&= p(1-p)\frac{d}{dp}(-\frac{1}{p}) = \frac{1-p}{p}
\end{aligned}
$$

Notice that

$$
\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=0}^{\infty} k^2(1-p)^{k-1}p \\
&= \sum_{k=0}^{\infty}[(k-1)^2 + 2(k-1) + 1](1-p)^{i-1}p \\
&= \sum_{k=0}^{\infty} k^2(1-p)^k p + 2\sum_{k=0}^{\infty} k(1-p)^2 p + 1
\end{aligned}
$$

We have $\mathbb{E}[X^2] = \frac{2-p}{p^2}$

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \frac{1-p}{p}
\end{aligned}
$$

<u>X follows Poisson distribtuion:</u>

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \\
&= e^{-\lambda}\sum_{k=1}^{\infty}\frac{k}{k!}\lambda^k \\
&= \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}e^{\lambda} = \lambda
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} \frac{1}{k!} \lambda^k e^{-k} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} (k-1+1) \frac{1}{(k-1)!} \lambda^{k-1} \\
&= \lambda e^{-\lambda} (\lambda \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k + \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k) \\
&= \lambda e^{-\lambda} (\lambda e^{\lambda} + e^{\lambda}) \\
&= \lambda(\lambda + 1)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \lambda^2 + \lambda - \lambda^2 = \lambda
\end{aligned}
$$

$\square$