DTU Compute

02425 Diffusions and SDEs
September 12, 2022
UHT/uht

Exercise 2: Probability measures

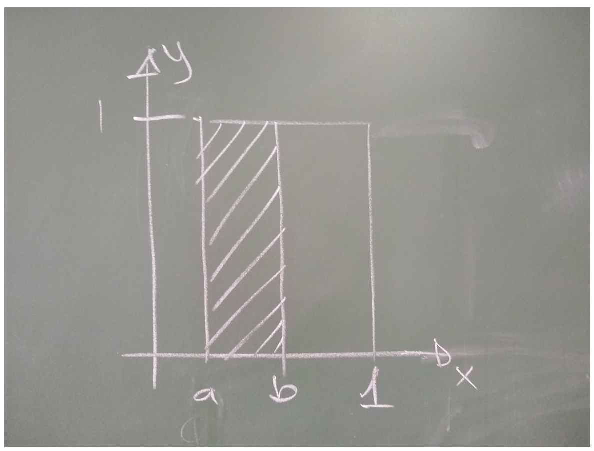
Consider the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with $\Omega = [0,1]^2$, $\mathcal{F}$ the usual Borel-algebra on Ω, and $\mathbf{P}$ the Lebesgue measure, i.e. area. For $\omega = (x,y) \in \Omega$, define $X(\omega) = x$, $Y(\omega) = y$, and $Z(\omega) = x + y$.

Question 1 **Independence and marginal distributions of X and Y:** Verify that X and Y are independent, and that each are uniformly distributed on $[0,1]$.

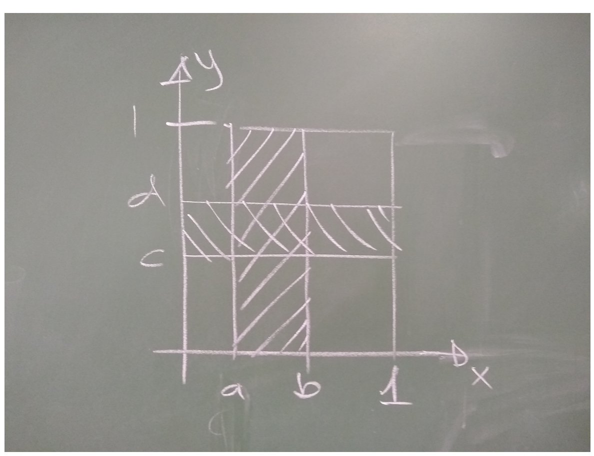
Solution: The following figure shows the sample space and the event $X \in [a,b]$. We see that the probability of this event is the area of the rectangle, i.e. $b - a$. With $a = 0$, we find

$$\mathbb{P}(X \leq b) = b$$

i.e., a uniform distribution.



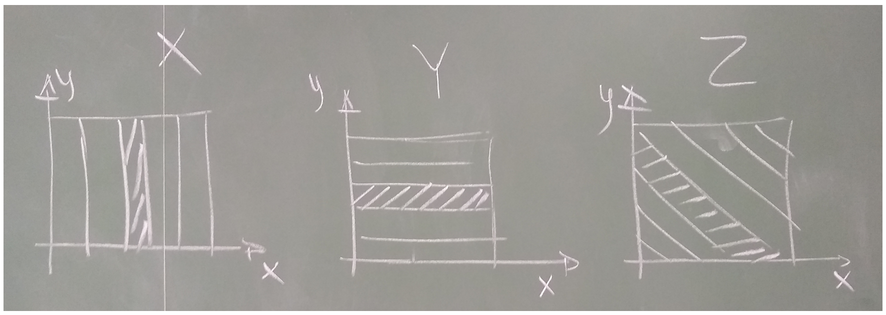
The next figure adds the event $Y \in [c,d]$. We see that $\mathbb{P}(X \in [a,b], Y \in [c,d]) = (b-a)(d-c) = \mathbb{P}(X \in [a,b])\mathbb{P}(Y \in [c,d])$, which defines independence.



Question 2 **Level sets and σ-algebras:** Sketch level sets (contour lines) for X, Y, and Z. *Note:* A level set for a function is a subset of the domain where the function attains one specific value,

such as $X^{-1}(1/2) = \{(x,y) : x = 1/2\}$. Show typical elements in the σ-algebras $\sigma(X)$, $\sigma(Y)$ and $\sigma(Z)$.

Solution: The following figure shows typical level sets for X, Y and Z, along typical elements in the σ-algebras: Level sets for X are lines where $X(\omega) = x$ are constant, and typical elements in the σ-algebra are regions between the level sets ("cylinder sets"). Similarly for Y and Z.



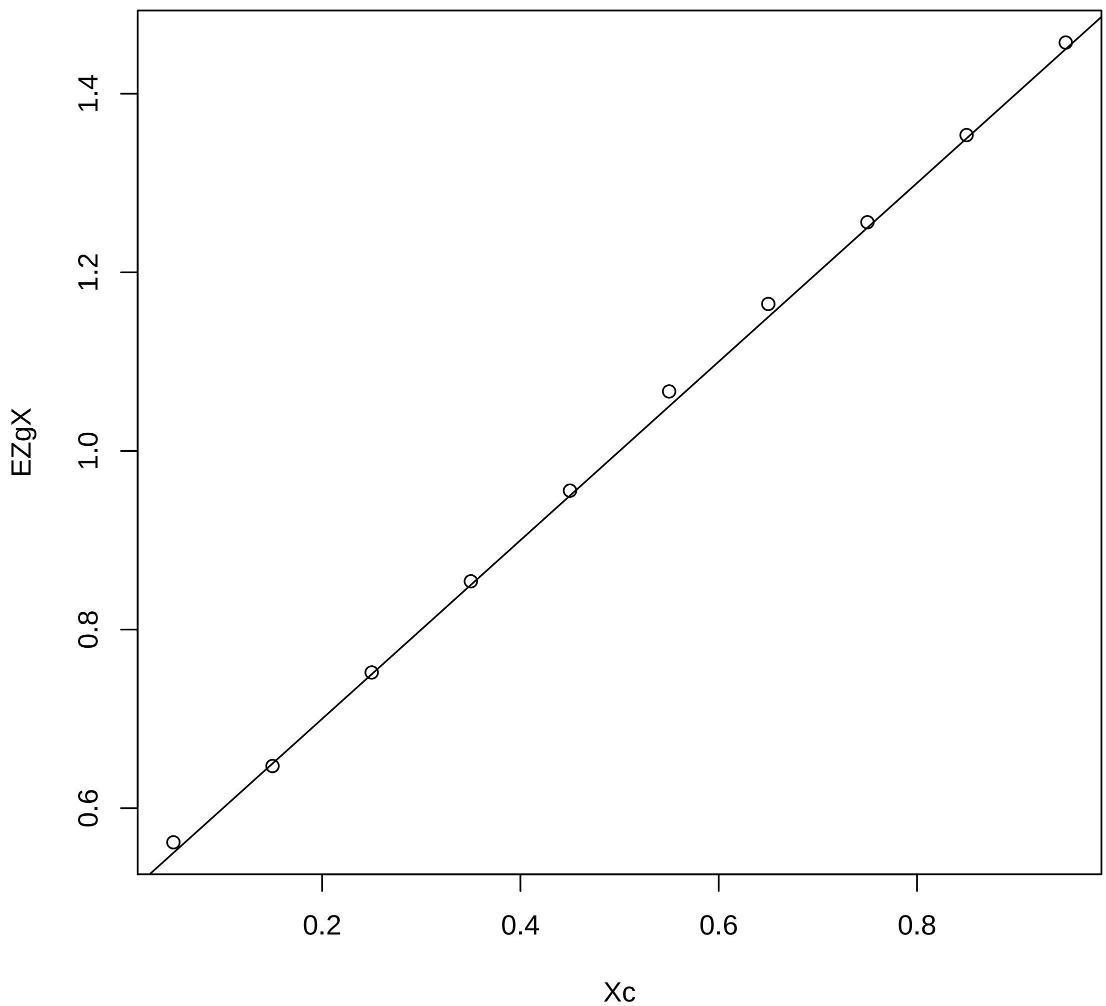
Question 3 Monte Carlo simulation and conditional expectations of Z given X: Simulate a larger number, for example $N = 10,000$, of realizations of X, Y and Z. Binning the realized values of X into bins $[0, 1/n, 2/n, \ldots, 1]$, estimate empirically the conditional expectations of Z given X

$$\mathbf{E}\left\{Z|X \in \left[\frac{i-1}{n}, \frac{i}{n}\right]\right\}$$

for $n = 10$ and $i = 1, \ldots, n$. Plot the result against i/n.

Solution: The following code does the simulation, the binning and the conditional expectations. We include also the analytical result from the next question.

```
N <- 1e4

X <- runif(N)
Y <- runif(N)
Z <- X+Y

dat <- data.frame(X=X,Y=Y,Z=Z)

Xbreaks <- seq(0,1,0.1)
Xc <- Xbreaks[-1] - 0.5*diff(Xbreaks)

EZgX <- tapply(Z,cut(X,breaks=Xbreaks),mean)

plot(Xc,EZgX)
plot(function(x)x+0.5,add=TRUE,from=0,to=1)
```



Question 4 **Conditional expectations of Z given X, analytically:** Find a function $g : \mathbf{R} \mapsto \mathbf{R}$ such that $\mathbf{E}\{Z|X\} = g(X)$. Use inspiration (!) from the Monte Carlo experiments and elementary arguments. Sketch contour lines of $\mathbf{E}\{Z|X\}$. Verify that this "candidate" conditional expectation satisfies the defining property

$$\int_H g(X(\omega))\ d\mathbf{P}(\omega) = \int_H Z(\omega)\ d\mathbf{P}(\omega)$$

for any $H \in \sigma(X)$.

Solution: We get:

$$\begin{aligned}
\mathbf{E}\{Z|X\} &= \mathbf{E}\{X+Y|X\} \quad \text{(definition of } Z\text{)} \\
&= \mathbf{E}\{X|X\} + \mathbf{E}\{Y|X\} \quad \text{(linearity of conditional expectations)} \\
&= X + \mathbf{E}\{Y\} \quad (X \text{ is } X\text{-measurable; } Y \text{ and } X \text{ are independent)} \\
&= X + \frac{1}{2} \quad (Y \text{ is uniformly distributed)}
\end{aligned}$$

so $g(x) = x + \frac{1}{2}$.

The contour lines of $\mathbf{E}\{Z|X\}$ agree with those of X, but the levels are different (the difference being

$\frac{1}{2}$).

To verify the definition of conditional expectation, we take H to be the cylinder set

$$H = \{\omega = (x, y) : a \le x \le b, 0 \le y \le 1\}.$$

We get

$$\int_H g(X(\omega))\ d\mathbf{P}(\omega) = \int_a^b \int_0^1 x + \frac{1}{2}\ dy\ dx = \frac{1}{2}(b^2 - a^2) + \frac{b-a}{2}.$$

Meanwhile, we get

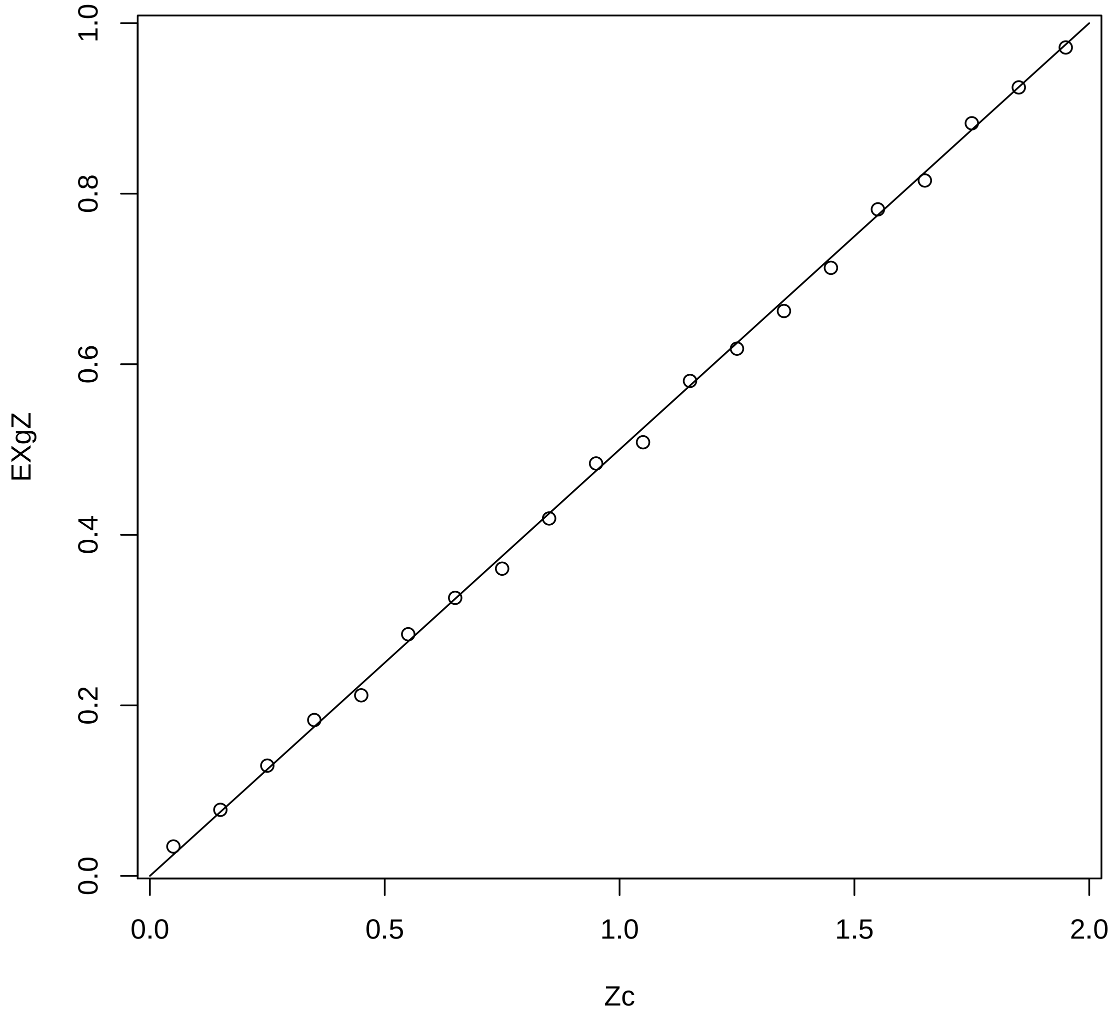
$$\int_H Z(\omega)\ d\mathbf{P}(\omega) = \int_a^b \int_0^1 x + y\ dy\ dx = +\frac{1}{2}(b^2 - a^2) + \frac{b-a}{2}$$

which verifies the definition.

Question 5 **Conditional expectations of X given Z:** Repeat the two previous questions for $\mathbf{E}\{X|Z\}$: First estimate the function $h : \mathbf{R} \mapsto \mathbf{R}$ such that $h(Z) = \mathbf{E}\{X|Z\}$ from the Monte Carlo experiments. Then find that function and verify that it satisfies its defining properties.

Solution: The following code does the simulation. We include the analytical result $\mathbf{E}\{X|Z\}$, which we find in the following.

```
Zbreaks <- seq(0,2,0.1)
Zc <- Zbreaks[-1] - 0.5*diff(Zbreaks)

cZ <- cut(Z,breaks=Zbreaks)

EXgZ <- tapply(X,cZ,mean)
plot(Zc,EXgZ)
plot(function(z)0.5*z,add=TRUE,from=0,to=2)
```



The result $\mathbf{E}\{X|Z\} = \frac{1}{2}Z$ is most easily obtained with an appeal to symmetry: We must have

$$\mathbf{E}\{X|Z\} + \mathbf{E}\{Y|Z\} = \mathbf{E}\{Z|Z| = Z$$

and

$$\mathbf{E}\{X|Z\} = \mathbf{E}\{Y|Z\}$$

Thus $h(z) = z/2$.

To verify that this function satisfies the defining properties, we first define $\hat{X} = \mathbf{E}\{X|Z\}$. We need a "typical element" $H \in \sigma(Z)$; we have previously sketched such an element. For simplicity, we consider a "thin" set where $Z \in [z, z + dz]$. There are two cases:

1. $z < 1$. Then the area of the set is $z\ dz$. We get $\hat{X} = z/2$ and $\mathbf{E}\{\hat{X}\mathbf{1}_H\} = z^2/2\ dz$. For comparison, we get $\mathbf{E}\{X\mathbf{1}_H\} = \int_0^z x\ dx\ dz = z^2/2\ dz$.

2. $z > 1$. Then the area of the set is $(2 - z)\ dz$. We get $\mathbf{E}\{\hat{X}|Z\} = \frac{z}{2}(2 - z)\ dz$. For comparison, we get $\mathbf{E}\{X\mathbf{1}_H\} = \int_{z-1}^1 x\ dx\ dz = \frac{1}{2}(1 - (z - 1)^2)\ dz = \frac{1}{2}(2z - z^2)\ dz = \frac{z}{2}(2 - z)\ dz$.

In either case, we see that $Z/2$ satisfies the defining property of $\mathbf{E}\{X|Z\}$.

Question 6 The law of total expectation (The simple Tower property): Verify from the Monte Carlo experiments that

$$\mathbf{E}\{\mathbf{E}\{X|Z\}\} = \mathbf{E}\{X\}$$

Solution: The verification of the can be done in two ways: We can use that $\hat{X} = Z/2$ and verify that $EX = \mathbf{E}\hat{X}$:

```
## Based on analytical conditioning:
mean(X)

## [1] 0.5004579

mean(Z/2)

## [1] 0.5037824
```

Alternatively, we can do this emprirically, where we condition not on true Z, but on the bin in which Z lies:

```
## Based on binned Z's:
mean(X)

## [1] 0.5004579

pZ <- table(cZ) / length(Z)
sum(EXgZ*pZ)

## [1] 0.5004579
```

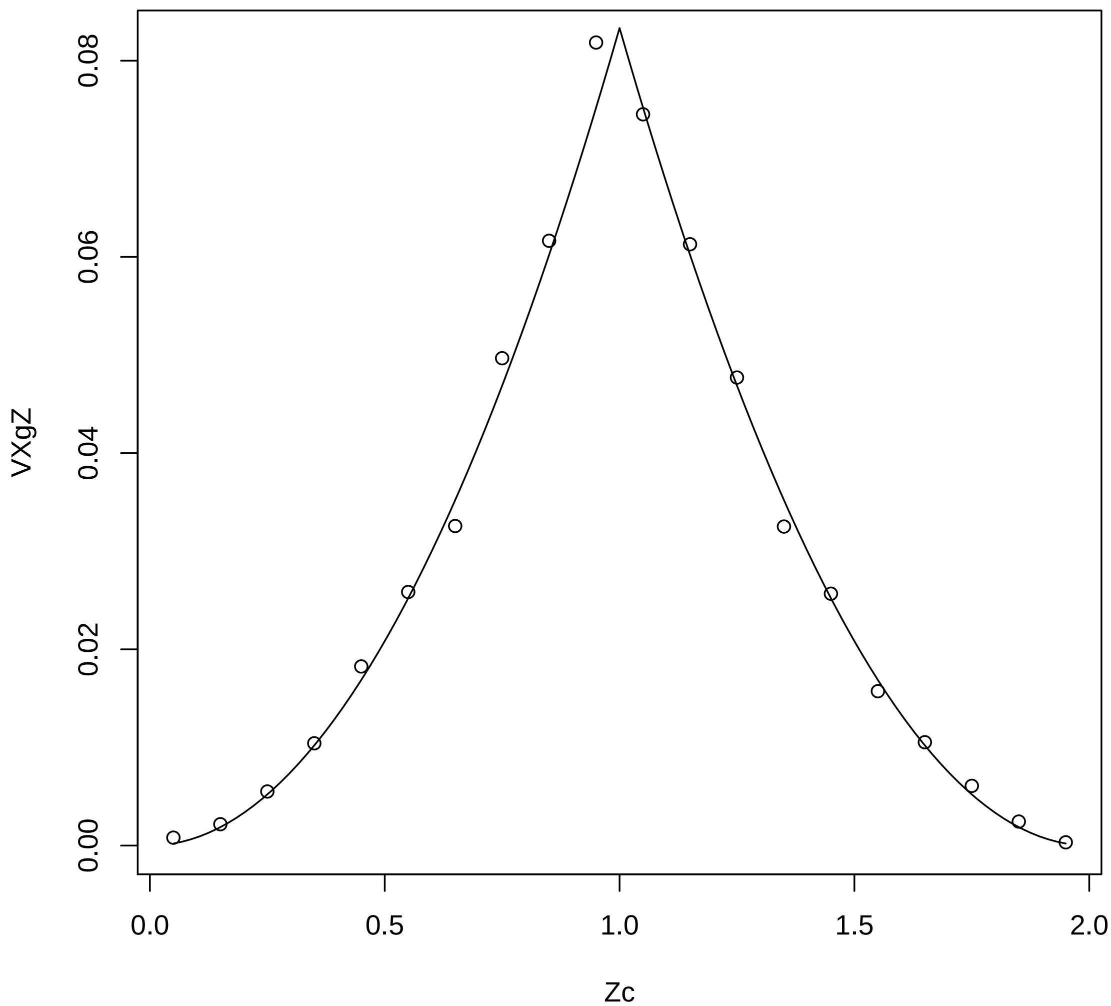
In the first case, we get a slight discrepancy due to the random Y-values. In the second case, we get exact equality.

Question 7 Variance decomposition/Law of total variance: Find an analytical expression for $\mathbf{V}\{X|Z\}$ and verify it against the Monte Carlo simulations. Then verify from the Monte Carlo simulations the variance decomposition formula

$$\mathbf{V}\{X\} = \mathbf{EV}\{X|Z\} + \mathbf{VE}\{X|Z\}$$

Solution: If $Z < 1$, then conditional on Z, X is uniformly distributed on $[0, Z]$ and therefore has conditional variance $Z^2/12$. If $Z > 1$, then X is uniformly distributed on $[Z-1, 1]$ and therefore has conditional variance $(2-Z)^2/12$. The following code verifies this expression:

```
## Conditional variance
VXgZ <- tapply(X,cZ,var)

plot(Zc,VXgZ)
VXgzanalytical <- function(z) pmin(z,2-z)^2/12
curve(VXgzanalytical,add=TRUE)
```



The verification can again be done purely empirical, or using the analytical expressions:

```
## Variance decomposition, analytically:
print(var(X))

## [1] 0.0836867

print(mean(VXgzanalytical(Z)) + var(Z/2))

## [1] 0.08340128

## Variance decomposition, purely empirical:
VXgZ[is.na(pZ)] <- 0
EXgZ[is.na(pZ)] <- 0
pZ[is.na(pZ)] <- 0

print(sum(pZ*VXgZ) + sum(pZ*(EXgZ-sum(pZ*EXgZ))^2))

## [1] 0.0837349
```