Problem (13.7)

In this problem we consider the Gaussian (or normal) distribution $N(\xi, \sigma^2)$, where $\xi \in \mathbb{R}$ and $\sigma^2 > 0$, i.e. the probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with density

$$f_{\varepsilon,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\xi)^2\right), \quad (x \in \mathbb{R})$$

with respect to the Lebesgue measure λ. We start by focusing on the case where $\xi = 0$, and $\sigma^2 = 1$, and we consider accordingly a random variable X on a probability space $(\Omega, \mathcal{F}, P)$, which is Gaussian distributed with parameters $(0, 1)$.

(a) Use Dominated Convergence and integration by parts to verify the identity:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^p \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) = \frac{p-1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{p-2} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)$$

for all p in $\mathbb{N}$, such that $p \geq 2$.

(b) Use an induction argument to verify that for any p in $\mathbb{N}$, it holds that

$$\mathbb{E}\left[X^p\right] = \begin{cases} (p-1)(p-3)\cdots 1, & \text{if } p \text{ is even} \\ 0, & \text{if } p \text{ is odd} \end{cases}$$

(c) Show that if Y is a random variable on $(\Omega, \mathcal{F}, P)$, which is Gaussian distributed with parameters (ξ, σ^2), then the random variable $(\sigma^2)^{-1/2}(Y - \xi)$ is Gaussian distributed with parameters $(0, 1)$.

(d) Assume that Y is a random variable on $(\Omega, \mathcal{F}, P)$, which is Gaussian distributed with parameters (ξ, σ^2). Argue then that Y has finite p' th moment for any p in $\mathbb{N}$, and calculate $\mathbb{E}[Y], \mathrm{Var}[Y]$ and the moments $\mathbb{E}\left[Y^3\right]$ and $\mathbb{E}\left[Y^4\right]$.

Solution

Part (a)

Since exp is a positive operation, we use monotone convergence to rewrite the integral:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^p \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) = \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \int_{-n}^{n} x^p \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)$$

We now use integration by parts:

$$\int u dv = uv - \int v du$$

We identify:

$$u = x^{p-1} \Rightarrow du = (p-1)x^{p-2}$$

$$dv = xe^{-\frac{1}{2}x^2} \Rightarrow v = -e^{-\frac{1}{2}x^2}$$

So we insert this into the formula:

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \left(\left[-x^{p-1} e^{-\frac{1}{2}x^2} \right]_{-n}^{n} - \int_{-n}^{n} -(p-1)x^{p-2} e^{-\frac{1}{2}x^2} \right) \\
&= \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \left(\left[-x^{p-1} e^{-\frac{1}{2}x^2} \right]_{-n}^{n} + \int_{-n}^{n} (p-1)x^{p-2} e^{-\frac{1}{2}x^2} \right) \\
&= \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \left(\int_{-n}^{n} (p-1)x^{p-2} e^{-\frac{1}{2}x^2} \right) \\
&= \frac{p-1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{p-2} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)
\end{aligned}$$

Part (b)

Case 1: $p = 1$
We use our result in a) to rewrite the expectation of the $p'th$ moment to:

$$\mathbb{E}[X^1] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) = 0$$

Or just because X has mean 0.
Case 2: $p = 2$ We use our result in a):

$$\begin{aligned}
\mathbb{E}[X^2] &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^0 \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)
\end{aligned}$$

We notice that this is just the density of a $N(0,1)$ distribution, and we know that it integrates to 1. Therefore:

$$\mathbb{E}[X^2] = 1$$

Inductive step: Assume that the formula holds for some $p = k$. We now need to show that it holds for $k+1$.
Case 1: k is even If k is even we have:

$$\mathbb{E}[X^k] = (k-1)(k-3)\cdots 1$$

And we must show that $\mathbb{E}[X^{k+1}] = 0$ since $k+1$ is odd.
In a) we showed that if $p \geq 2$, we have

$$\mathbb{E}[X^p] = (p-1)\mathbb{E}[X^{p-2}]$$

By this recursive structure, if we keep doing it we will end up where $\mathbb{E}[X^{p-n}] = \mathbb{E}[X^1] = 0$ so therefore it holds for $p = k+1$ odd.
Case 2: k is odd If k is odd we have $k+1$ is even. In a we showed that:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^p \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) = \frac{p-1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{p-2} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)$$

for all p in $\mathbb{N}$, such that $p \geq 2$, and thus:

$$\int_{\mathbb{R}} x^p \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x) = (p-1) \int_{\mathbb{R}} x^{p-2} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)$$

So if we put in $k+1$ in we get:

$$\mathbb{E}[X^{k+1}] = \frac{k}{\sqrt{2\pi}} \int_{\mathbb{R}} x^{k-1} \exp\left(-\frac{1}{2}x^2\right) \lambda(\mathrm{d}x)$$

Which is $(k+1)-1(E[X^{k-1}])$ and so on. So by the recursive nature of the moments, we have that it holds for $p = k+1$ if $k+1$ is even.

Part (c)

Let $Y \sim N(\xi, \sigma^2)$, so the density function is given by:

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y-\xi)^2\right), \quad (y \in \mathbb{R})$$

Now we consider random variable $Z = (\sigma^2)^{-0.5}(Y - \xi) = \frac{Y-\xi}{\sqrt{\sigma^2}}$
So the inverse is equal to $Y = \xi + \sqrt{\sigma^2}Z$ with $\frac{dY}{dZ} = \sqrt{\sigma^2}$
By 13.3.7 we have

$$f_Z(z) = f_Y(\xi + \sqrt{\sigma^2}z) \left| \frac{d}{dz}(\xi + \sqrt{\sigma^2}z) \right|$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}((\xi + \sqrt{\sigma^2}z) - \xi)^2\right) \sqrt{\sigma^2}$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}((\xi + \sqrt{\sigma^2}z) - \xi)^2\right)$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right)$$

Which is the desired result.

Part (d)

$$\mathbb{E}[Y] = \xi$$
$$\mathrm{Var}[Y] = \sigma^2$$
$$\mathbb{E}[Y^3] = \xi^3 + 3\xi\sigma^2$$
$$\mathbb{E}[Y^4] = 3\sigma^2\xi^3 + 6\xi^2\sigma^2 + \xi^4$$

Problem (13.10)

Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, P)$, and assume that $X \in \mathcal{L}^2(P)$.

(a) Show that $\mathrm{Var[X]} = \mathrm{E[X(X-1)]} - \mathbb{E}[\mathrm{X}](\mathbb{E}[\mathrm{X}]-1)$.

(b) Use (a) to calculate $\mathrm{Var}[X]$, when X has a binomial distribution.

(c) Use (a) to calculate Var[X], when X has a Poisson distribution.

Solution

Part (a)

Recall that

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

And express $X^2 = X(X-1) + X$ Then

$$\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X]$$

We now input this into our equation for the variance:

$$\mathrm{Var}[X] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2$$

$$= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]\mathbb{E}[X]$$

$$= \mathbb{E}[\mathrm{X(X-1)}] - \mathbb{E}[\mathrm{X}](\mathbb{E}[\mathrm{X}]-1)$$

Part (b)

We have $\mathbb{E}[X] = np$ and $\mathbb{E}[X(X-1)] = n(n-1)^2$
So by (a) we have:

$$\mathrm{Var}[X] = \mathbb{E}[\mathrm{X(X-1)}] - \mathbb{E}[\mathrm{X}](\mathbb{E}[\mathrm{X}]-1)$$

$$= n(n-1)p^2 - np(np-1)$$

$$= n(n-1)p^2 - np^2 + np$$

$$= np(1-p)$$

Part (c)

We have $\mathbb{E}[X] = \lambda$ and $\mathbb{E}[X(X-1)] = \lambda^2$
So by (a) we have:

$$\mathrm{Var}[X] = \mathbb{E}[\mathrm{X(X-1)}] - \mathbb{E}[\mathrm{X}](\mathbb{E}[\mathrm{X}]-1)$$

$$= \lambda^2 - \lambda(\lambda-1)$$

$$= \lambda^2 - (\lambda^2 - \lambda)$$

$$= \lambda$$

Problem (13.12)

Let $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ be two sequences of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Assume further that

$$P\left(|\mathrm{X}_n - \mathrm{Y}_n| > n^{-1}\right) \leq n^{-2} \quad \text{for all } n \text{ in } \mathbb{N}.$$

Show then that the following assertions are equivalent:

(a) There exists a random variable X on $(\Omega, \mathcal{F}, P)$ such that $X_n \to X$ P-n.o.

(b) There exists a random variable Y on $(\Omega, \mathcal{F}, P)$ such that $Y_n \to Y P$-n.o.

Show also that if (a) and (b) are satisfied, then $X = Y P$-n.o.

Solution

Part (a)

We have that it is enough to show one of the ways, since we have not assumed anything about the sequence $(X_n)_{n\in\mathbb{N}}$ that we have not assumed about $(Y_n)_{n\in\mathbb{N}}$. Therefore we only show (a)$\Rightarrow$(b). Note that:

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - Y_n| > n^{-}1) \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$$

So we use Borel-Cantelli I, where we set $F_n = \{|X_n - Y_1| > n^{-}1\}$ for all $n \in \mathbb{N}$, then $\mathbb{P}(\bigcup_{n\in\mathbb{N}} F_n) = 0$
So if $U = \bigcup_{n\in\mathbb{N}} F_n$, then

$$U^C = \{\omega \in \Omega : |X_n(\omega) - Y_n(\omega)| \leq n^{-}1\text{for n large enough}\}$$

We set $A = U^C \bigcap \{X_n \to X\}$ so for $\omega \in A$ there exists $N' \in \mathbb{N}$ such that

$$|X_n(\omega) - Y_n(\omega)| \leq \frac{1}{n} \forall n \geq N'$$

So $(X_n(\omega))_{n\in\mathbb{N}}$ converges, it is a cauchy sequence.
We now let $\epsilon > 0$, and choose $N'' \in \mathbb{N}$ such that:

$$\frac{1}{n} \leq \frac{\epsilon}{3}$$

and

$$|X_n(\omega) - X_m(\omega)| \leq \frac{\epsilon}{3} \forall n, m \geq N''$$

We now set $N = \max\{N', N''\}$. Then we have for $n, m \geq N$ it holds that:

$$|Y_n(\omega) - Y_m(\omega)| \leq |Y_n(\omega) - X_n(\omega)| + |X_n(\omega) - X_m(\omega)| + |X_m(\omega) - Y_m(\omega)| \leq 3\frac{\epsilon}{3} = \epsilon$$

So $(Y_n(\omega))_{n\in\mathbb{N}}$ is a cauchy sequence, so it converges. We set $Y = \lim_{n\to\infty} Y_n \mathbb{1}_A$. Then we have by 4.3.7(ii) that Y is a stochastic variable. Since we have:

$$\mathbb{P}(A^C) \leq \mathbb{P}(U) + \mathbb{P}(\{X_n \to X\}^C) = 0 + 0 = 0$$

So $Y_n \to Y$ p-a.e., so (a)$\Rightarrow$(b)
Now assume (a) and (b) both holds. Then we have ofr

$$B = \omega \in \{X_n \to X\} \cap \{Y_n \to Y\} \cap U^C$$

it holds for all $n \in \mathbb{N}$ that

$$|X(\omega) - Y(\omega)| \leq |X(\omega) - X_n(\omega)| + |X_n(\omega) - Y_n(\omega)| + |Y_n(\omega) - Y(\omega)|$$

and since the RHS tends to 0 as $n \to \infty$ then $X(\omega) = Y(\omega)$ and since $\mathbb{P}(B^C) = 0$ as per the same argument as above, then $X = Y$ p-a.e.

Problem (13.15)

Let X and Y be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$.

(a) Assume that $X, Y \in \mathcal{L}^2(P)$. Show then that if X and Y are independent, it holds that $\operatorname{Cov}[X,Y] = 0$, and that $\operatorname{Var}[X+Y] = \operatorname{Var}[X] + \operatorname{Var}[Y]$.

(b) Show that X and Y are independent, if and only if $\operatorname{Cov}[f(\mathrm{X}), g(\mathrm{Y})] = 0$ for all bounded Borel functions $f, g : \mathbb{R} \to \mathbb{R}$.

Hint for the "if-part": Let f and g be indicator functions.

Solution

Part (a)

It follows from corollary 13.5.6(i) that $XY \in \mathcal{L}^1(P)$, since X and Y are independent and has the second moment, and it further holds that:

$$\operatorname{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] = 0$$

and therefore it follows that:

$$\operatorname{Var}[X+Y] = \operatorname{Var}[X] + \operatorname{Var}[Y] - 2\operatorname{Cov}[X,Y] = \operatorname{Var}[X] + \operatorname{Var}[Y] - 2 \cdot 0 = \operatorname{Var}[X] + \operatorname{Var}[Y]$$

Part (b)

First we show $\Rightarrow$. Assume X and Y are independent and let $f, g : \mathbb{R} \to \mathbb{R}$ be bounded borel-functions, then it follows from remark 13.5.2(ii) that $f(X)$ and $g(Y)$ are independent, and since f, g are bounded, $f(X), g(Y) \in \mathcal{L}^2(\mathbb{P})$,and then by (a), $\operatorname{Cov}[f(X), g(Y)] = 0$.
We now show $\Leftarrow$. Assume $\operatorname{Cov}[f(X), g(Y)] = 0$ for all Borel-functions $f, g : \mathbf{R} \to \mathbf{R}$. Let $A, B \in \mathcal{F}$, then $\mathbf{1}_A, \mathbf{1}_B$ are bounded Borel-functions, then it follows that

$$\begin{aligned}
0 &= \operatorname{Cov}\left[\mathbf{1}_A(X), \mathbf{1}_B(Y)\right] = \mathbb{E}\left[\mathbf{1}_A(X)\mathbf{1}_B(Y)\right] - \mathbb{E}\left[\mathbf{1}_A(X)\right]\mathbb{E}\left[\mathbf{1}_B(Y)\right] \\
&= \mathbb{E}\left[\mathbf{1}_{\{X \in A, Y \in B\}}\right] - \mathbb{E}\left[\mathbf{1}_{\{X \in A]}\right]\mathbf{E}\left[\mathbf{1}_{[Y \in B]}\right] \\
&= \mathbb{P}(X \in A, Y \in B) - \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \\
&\implies \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

so X and Y are independent.

Problem (13.16)

Let X be a random variable defined on a probability space et $(\Omega, \mathcal{F}, P)$, and assume that the distribution of X is the uniform distribution on $[-1, 1]$, i.e. the measure with density

$$f_{\mathrm{X}}(t) = \frac{1}{2}\mathbf{1}_{[-1,1]}(t), \quad (t \in \mathbb{R})$$

with respect to Lebesgue measure λ on $\mathbb{R}$.

(a) Show that $\mathrm{Cov}\,[\mathrm{X}, \mathrm{X}^2] = 0$.

(b) Decide whether X and X^2 are independent.

Solution

Part (a)

We first see that $X, X^2 \in \mathcal{L}^2(P)$ since f_X is bounded. Since $x \mapsto x^2, x^3$ are borel functions, 13.3.6 entails that $x \mapsto x, x^3$ are uneven functions, and it follows that:

$$\mathrm{Cov}[X, X^2] = \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2]$$
$$= \frac{1}{2}\int_{-1}^{1} x^3 dx - \left(\frac{1}{2}\int_{-1}^{1} x dx\right)(\frac{1}{2}\int_{-1}^{1} x \cdot x^2 dx)$$
$$= 0 - 0 \cdot \left(\frac{1}{2}\int_{-1}^{1} x^3 dx\right) = 0$$

Part (b)

No. We see that the mean for $X \cdot X^2 = X^3$ does not split in one product since:

$$\mathbb{E}[X^3] = \int_{-1}^{1} x \cdot x^3 \,\mathrm{d}x = \int_{-1}^{1} x^4 \,\mathrm{d}x = \frac{2}{5}$$

but $\mathbb{E}[X^2] = 0$

Problem (13.17)

Let X be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$, and assume that $\mathbb{E}\left[X^2\right] < \infty$.

(a) Prove Chebychev's inequality:

$$P(|\mathrm{X} - \mathbb{E}[\mathrm{X}]| \geq \epsilon) \leq \frac{1}{\epsilon^2} \operatorname{Var}[\mathrm{X}]$$

for any positive number ϵ.

(b) Assume that $X_1, X_2, X_3, \ldots$ is a sequence of i.i.d. random variables with the same distribution as X. Prove then for any positive number ϵ that

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbf{E}\left| X \right]\right| > \epsilon\right) \xrightarrow[n\to\infty]{} 0$$

This result is a version of the so-called Weak Law of Large Numbers; compare with Proposition 13.6.1.

Solution

Part (a)

We first notice that $\Delta = \{(y, y) \in \mathbb{R}^2 : y \in R\}$ is a λ_2-null set as per remark 13.3.5 and since X and Y are independent, (X, Y) are absolutely continuous by example 13.5.7(A). So it follows from proposition 5.3.6(ii) that

$$\begin{aligned}
\mathbf{P}(X = Y) = \mathbb{P}((X, Y) \in \Delta) &= \int_{\mathbb{R}^2} \nVdash_\Delta \mathrm{d}\mathbb{P}_{(X,Y)} \\
&= \int_{\mathbb{R}^2} f_{(X,Y)}(x, y) \mathbf{1}_\Delta(x, y) \lambda_2(\,\mathrm{d}x,\ \mathrm{d}y) = 0
\end{aligned}$$

Part (b)

Since $\mathbf{P}(X = Y) = 0$, it follows that $\mathbf{P}(X < Y) + \mathbb{P}(X > Y) = 1$, and since X and Y are identically distributed, $\mathbb{P}_X = \mathbf{P}_Y$, since X and Y are independent, and $\mathbf{P}_{(X,Y)} = \mathbf{P}_X \otimes \mathbf{P}_Y$, then it holds that,

$$\begin{aligned}
\mathbf{P}(X < Y) &= \mathbb{P}_{(X,Y)}\left(\left\{(x, y) \in \mathbb{R}^2 : x < y\right\}\right) \\
&= (\mathbb{P}_X \otimes \mathbb{P}_Y)\left(\left\{(x, y) \in \mathbb{R}^2 : x < y\right\}\right) \\
&= (\mathbb{P}_Y \otimes \mathbb{P}_X)\left(\left\{(x, y) \in \mathbb{R}^2 : x < y\right\}\right) \\
&= \mathbb{P}_{(Y,X)}\left(\left\{(x, y) \in \mathbb{R}^2 : x < y\right\}\right) \\
&= \mathbb{P}(X > Y)
\end{aligned}$$

and therefore we have:

$$2\mathbb{P}(X < Y) = \mathbb{P}(X < Y) + \mathbb{P}(X > Y) = 1 \Longrightarrow \mathbb{P}(X < Y) = \mathbb{P}(X > Y) = \frac{1}{2}$$