EXERCISES 8

1.

Let $X(\omega)$ be a random variable with $P(X \geq 0) = 1$.

a) Claim: $E_P(X) = \int_0^\infty P(X > t)\,dt = \int_0^\infty P(X \geq t)\,dt$

Proof:

Direct proof:

$$E_P(X) = \int_0^\infty t \, dP_X(t) = \int_0^\infty \left(\int_0^\infty \mathbb{1}(s \overset{?}{<} t)\,ds \right) dP_X(t) \overset{(*)}{=} \int_0^\infty \left(\int_0^\infty \mathbb{1}(s \overset{?}{<} t)\,dP_X(t) \right)$$

$$ds = \int_0^\infty P(X \overset{?}{>} s)\,ds$$

where $\overset{?}{<} \in \{<, \leq\}$; we note that $t = \int_0^\infty \mathbb{1}(s < t)\,ds = \int_0^\infty \mathbb{1}(s \leq t)\,ds$ and we use Fubini in $(*)$. P_X is the distribution measure of X.

Hence the claim is true. □

b) Claim: $E_P(X^n) = n \int_0^\infty t^{n-1} P(X > t)\,dt$ for $n \in \mathbb{N}$.

Proof:

Direct proof:

$$E_P(X^n) = \int_0^\infty P(X^n > t)\,dt = \int_0^\infty P(X > t^{\frac{1}{n}})\,dt = \int_0^\infty P(X > s) \cdot n (s^n)^{1 - \frac{1}{n}}\,ds$$

$$= n \int_0^\infty s^{n-1} P(X > s)\,ds$$

where we use part a) and we perform a change of variable $s = t^{\frac{1}{n}}$; $ds = \frac{1}{n} t^{\frac{1}{n} - 1} dt = \frac{1}{n} (s^n)^{\frac{1}{n} - 1} dt$. Note that change of variable is applicable to measurable functions f ($\int f = \int (f \circ \varphi) \varphi'$).
Hence the claim is true. □

2.

Let $G(\omega) \sim \mathcal{N}(0,1)$ be a standard Gaussian random variable with probability density

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \; x \in \mathbb{R}.$$

Since we know that $E_P(\exp(\frac{1}{2}\lambda G^2)) < \infty$ $\forall \lambda < 1$, and when $0 < \lambda < 1$, for any polynomial $p(x)$, there are constants C_1, C_2 such that

$$|p(x)| \leq C_1 + C_2 \exp(\tfrac{1}{2}\lambda x^2)$$

which implies that $E_P(|G|^p) < \infty$ and $G \in L^p(P)$ for all exponents $p > 0$. Note also that the standard Gaussian distribution is symmetric around the origin, with $\phi(x) = \phi(-x)$.

a) Claim: $\forall n \in \mathbb{N}_0: E_P(G^{2n+1}) = 0$ (all the odd moments vanish).

Proof:

Direct proof:

For $n \in \mathbb{N}_0$,

$$E_P(G^{2n+1}) = \int_{-\infty}^{\infty} x^{2n+1}\phi(x)dx = \int_{\infty}^{-\infty}(-y)^{2n+1}\phi(-y)(-dy) = (-1)^{2n+1}\int_{-\infty}^{\infty} y^{2n+1}$$

$$\phi(y)dy = -E_P(G^{2n+1})$$

$$\Rightarrow E_P(G^{2n+1}) = 0$$

where we make the change of variable $y = -x$, $dy = -dx$ and use the symmetry $\phi(-y) = \phi(y)$.

Hence the claim is true. $\square$

b) Let us compute $E_P(G^2)$.

By c)-part, $E_P(G^2) = 1$.

c) Let us use induction to compute the even moments of the standard Gaussian $E_P(G^{2n})$ for $n \in \mathbb{N}$.

Let us note that $E_P(G^0) = E_P(1) = 1$.

By integration by parts:

$$E_P(G^{2n}) = \int_{-\infty}^{\infty} x^{2n}\phi(x)dx = \left[\frac{x^{2n+1}}{2n+1}\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\right]_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{x^{2n+1}}{2n+1}\cdot(-\tfrac{1}{2}\cdot 2x)\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx$$

$$= 0 - 0 + \frac{1}{2n+1}\int_{-\infty}^{\infty} x^{2(n+1)}\phi(x)dx = \frac{1}{2(n+1)-1}E_P(G^{2(n+1)}).$$

Above we note $\lim_{x\to\pm\infty} x^{2n+1}e^{-\frac{1}{2}x^2} = 0$, $\partial_x\phi(x) = -x\phi(x)$.

By induction we get

$$E_P(G^{2n}) = (2n-1)E_P(G^{2(n-1)}) = (2n-1)(2(n-1)-1)E_P(G^{2(n-2)})$$

$$= (2n-1)\cdot(2(n-1)-1)\cdot\ldots\cdot(2(n-(n-1))-1)\cdot E_P(G^{2(n-n)}) = (2n-1)$$

$$\cdot(2n-3)\cdot\ldots\cdot 1\cdot 1 = (2n-1)!! = \prod_{\substack{k<2n\\ k \text{ odd}}} k$$

3.

For $t\in\mathbb{R}$ let us compute the expectations

a) $E_P(G\,\mathbb{1}(G>t))$

b) $E_P(G\,\mathbb{1}(G\le t))$

c) $E_P(G^2\,\mathbb{1}(G>t))$

d) $E_P(G^2\,\mathbb{1}(G\le t))$

e) $E_P(G^3\,\mathbb{1}(G>t))$

f) $E_P(G\,\mathbb{1}(G\le t))$

Let us first show that the Gaussian integration by parts formula $E_P(f(G)G) = E_P(f'(G))$ holds with $f(x) = \mathbb{1}(x>t)$. In this case $f'(x) = \delta_t(x) = \delta_0(x-t)$ is not a function but a generalised function (a distribution in analysis language), the Dirac-delta function at t, with the defining property

$$g(t) = \int_{\mathbb{R}} g(x)\delta_t(x)dx = \int_{\mathbb{R}} g(x)\delta_0(x-t)dx = \int_{\mathbb{R}} g(y+t)\delta_0(y)dy$$

for any continuous test function g with compact support. From the probabilistic point of view the measure $\mu(dx) = \delta_t(x)dx$ is simply the probability measure of a deterministic random variable concentrated in the singleton $\{t\}$.

Claim: $E_P(f(G)G) = E_P(f'(G))$ for $f(x) = \mathbb{1}(x>t)$ where $f'(x) = \delta_t(x)$.

Proof:

Direct proof:

Let us define

$$f_n(x) := \min\{ n(x-t)^+ ; 1\} \quad \forall n \in \mathbb{N}$$

where $(x-t)^+ = \begin{cases} 0, & x-t \le 0 \\ x-t, & x-t \ge 0 \end{cases}$

Obviously the sequence (f_n) approximates the indicator $f(x) = \mathbb{1}(x>t)$ so that

$$0 \le f_n(x) \le f(x) \le 1 \quad \forall x \in \mathbb{R}$$

and also f_n is piecewise linear and continuous with derivative $f_n'(x) = n\,\mathbb{1}(t < x < t+\frac{1}{n})$, and

$$f_n(x) \uparrow f(x) \quad \forall x \in \mathbb{R}.$$

By the Gaussian integration by parts (see Gasbarra: Lecture notes in Probability theory fall semester 2015: Chapter 7) applied to f_n ($E_P(|f_n'(G)|) \le n\, E_P(|G|) < \infty$)

$$E_P(f(G)G) \overset{(*)}{=} \lim_{n\to\infty} E_P(f_n(G)G) = \lim_{n\to\infty} E_P(f_n'(G)) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n'(x)\,\phi(x)\,dx = \lim_{n\to\infty} n \int_t^{t+\frac{1}{n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} = \int_{-\infty}^{\infty} \phi(x)\,\delta_t(x)\,dx$$

$$=: E_P(f'(G))$$

where in $(*)$ we use Dominated convergence theorem ($|f_n(G)G| \le |G|$, $E_P(|G|) < \infty$) and we note that

$$\left| n \int_t^{t+\frac{1}{n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \right| = \left| n \int_t^{t+\frac{1}{n}} \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}x^2} - e^{-\frac{1}{2}t^2}\right) dx \right|$$

$$\le n \int_t^{t+\frac{1}{n}} \frac{1}{\sqrt{2\pi}} \left| e^{-\frac{1}{2}x^2} - e^{-\frac{1}{2}t^2} \right| dx \le n \cdot \frac{1}{n} \cdot \frac{1}{\sqrt{2\pi}}\,\varepsilon < \varepsilon$$

as $n \in \mathbb{N}$ is so large that $x, y \in [t, t+1] \wedge |x-y| < \frac{1}{n} \Rightarrow |e^{-\frac{1}{2}x^2} - e^{-\frac{1}{2}y^2}$

$\leq \varepsilon$. This proves $\lim_{n\to\infty} n \int_t^{t+\frac{1}{n}} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$.

Hence the claim is true. $\square$

a) $E_P(G\,\mathbb{1}(G>t)) = E_P(f(G)G) = E_P(f'(G)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$ by above.

b) $E_P(G\,\mathbb{1}(G\leq t)) = E_P(G) - E_P(G\,\mathbb{1}(G>t)) = 0 - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$

where we note $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E_P(G)=0$, $E_P(|G|)<\infty$.

c) We see that the above proof can be generalized to

$$E_P(g(G)f(G)G) = \lim_{n\to\infty} E_P(g(G)f_n(G)G) = \lim_{n\to\infty}\Big(E_P(g'(G)f_n(G))$$
$$+E_P(g(G)f_n'(G))\Big) = E_P(g'(G)f(G)) + \lim_{n\to\infty} n\int_t^{t+\frac{1}{n}} g(x)\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}dx$$
$$= E_P(g'(G)f(G)) + g(t)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} =: E_P(g'(G)f(G)) + E_P(g(G)f'(G)).$$

where $g:\mathbb{R}\to\mathbb{R}$ is a C^1-function such that $E_P(|g(G)G|)<\infty$, $E_P(|g'(G)|)<\infty$

Thus

$$E_P(G^2\,\mathbb{1}(G>t)) = E_P(f(G)) + E_P(Gf'(G)) = P(G>t) + t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$
$$= \int_t^\infty \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx + t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$

d) $E_P(G^2\,\mathbb{1}(G\leq t)) = E_P(G^2) - E_P(G^2\,\mathbb{1}(G>t)) = 1 - \int_t^\infty \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx - t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$
$$= \int_{-\infty}^t \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx - t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$

where $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E(G^2)=1$.

e) $E_P(G^3\,\mathbb{1}(G>t)) = E_P(2Gf(G)) + E_P(G^2 f'(G)) = 2\cdot\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} + t^2\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$

f) $E_P(G^3\,\mathbb{1}(G\leq t)) = E_P(G^3) - E_P(G^3\,\mathbb{1}(G>t)) = 0 - (2+t^2)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} = -(2+t^2)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$

where $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E(G^3)=0$

5.

(χ_n^2 and Dirichlet distributions)

$\leq \varepsilon$. This proves $\lim_{n\to\infty} n \int_t^{t+\frac{1}{n}} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$.

Hence the claim is true. $\square$

a) $E_P(G\,\mathbb{1}(G>t)) = E_P(f(G)G) = E_P(f'(G)) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$ by above.

b) $E_P(G\,\mathbb{1}(G\leq t)) = E_P(G) - E_P(G\,\mathbb{1}(G>t)) = 0 - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$

where we note $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E_P(G)=0$, $E_P(|G|)<\infty$.

c) We see that the above proof can be generalized to

$$E_P(g(G)f(G)G) = \lim_{n\to\infty} E_P(g(G)f_n(G)G) = \lim_{n\to\infty}\Big(E_P(g'(G)f_n(G))$$
$$+ E_P(g(G)f_n'(G))\Big) = E_P(g'(G)f(G)) + \lim_{n\to\infty} n \int_t^{t+\frac{1}{n}} g(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$
$$= E_P(g'(G)f(G)) + g(t)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} =: E_P(g'(G)f(G)) + E_P(g(G)f'(G)).$$

where $g: \mathbb{R}\to\mathbb{R}$ is a C^1-function such that $E_P(|g(G)G|)<\infty$, $E_P(|g'(G)|)<\infty$

Thus

$$E_P(G^2\,\mathbb{1}(G>t)) = E_P(f(G)) + E_P(G f'(G)) = P(G>t) + t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$
$$= \int_t^\infty \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx + t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$$

d) $E_P(G^2\,\mathbb{1}(G\leq t)) = E_P(G^2) - E_P(G^2\,\mathbb{1}(G>t)) = 1 - \int_t^\infty \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx - t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$
$= \int_{-\infty}^t \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx - t\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$

where $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E(G^2)=1$.

e) $E_P(G^3\,\mathbb{1}(G>t)) = E_P(2Gf(G)) + E_P(G^2 f'(G)) = 2\cdot\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} + t^2\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$

f) $E_P(G^3\,\mathbb{1}(G\leq t)) = E_P(G^3) - E_P(G^3\,\mathbb{1}(G>t)) = 0 - (2+t^2)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} = -(2+t^2)\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}$

where $\mathbb{1}(G\leq t) + \mathbb{1}(G>t) = 1$, $E(G^3)=0$

5.

(χ_n^2 and Dirichlet distributions)

Let $G(\omega)=(G_1(\omega);\dots;G_n(\omega))$ be independent and identically distributed standard Gaussian random variables, each with probability density $\phi(x)$ on $\mathbb{R}$.

Let

$$X(\omega)=G_1(\omega)^2+\dots+G_n(\omega)^2, \text{ and } \pi_k(\omega)=\frac{G_k(\omega)^2}{X(\omega)},\ 1\le k\le n$$

Note that $\pi_k(\omega)\in[0,1]$ and $\sum_{k=1}^{n}\pi_k(\omega)=1$ with probability 1. This means that (with probability 1) for each ω, the random vector $\pi(\omega)=(\pi_1(\omega);\dots;\pi_n(\omega))$ belongs to the simplex

$$\Delta_n=\{p=(p_1;\dots;p_n)\in[0,1]^n \mid p_1+\dots+p_n=1\}$$

and the vector determines a (random) probability distribution on the discrete set $\{1,\dots,n\}$. Note also that π is determined by $(n-1)$ coordinates, since $p_1=1-(p_2+\dots+p_n)$.

By the way, the distribution of X is called chi-square distribution with n degrees of freedom and it is denoted by χ_n^2, while the distribution of the random probability vector $(\pi_1;\dots;\pi_n)$ is a special case of the Dirichlet distribution, which is used to model random discrete probabilities.

Claim: X and π are independent.

Proof:

Direct proof: It is enough to prove that X and $(\pi_2;\dots;\pi_n)$ are independent, as π is a Borel map of $(\pi_2;\dots;\pi_n)$. Let us define $f:\mathbb{R}_+^n\to\mathbb{R}_+\times[0,1]^{n-1}$

$$g=(g_1;\dots;g_n)\mapsto f(g)=\Big(\sum_{k=1}^{n}g_k^2;\ \frac{g_2^2}{\sum_{k=1}^{n}g_k^2};\dots;\frac{g_n^2}{\sum_{k=1}^{n}g_k^2}\Big)$$

f is a bijection. Let us choose a measurable bounded test function $h:\mathbb{R}^n\to\mathbb{R}$ and calculate $(P=P_{G_1}\otimes\dots\otimes P_{G_n})$

$$E_P(h(X,(\pi_2;\dots;\pi_n)))=\int_{\mathbb{R}^n} h\Big(x_1^2+\dots+x_n^2;\ \frac{x_2^2}{\sum_{k=1}^{n}x_k^2};\dots;\frac{x_n^2}{\sum_{k=1}^{n}x_k^2}\Big)$$

$$\phi(x_1)\cdots\phi(x_n)dx_1\cdots dx_n = \int_{\mathbb{R}^n} h\left(x_1^2+\cdots+x_n^2;\frac{x_2^2}{\sum_{k=1}^n x_k^2};\cdots;\frac{x_n^2}{\sum_{k=1}^n x_k^2}\right)\left(\frac{1}{\sqrt{2\pi}}\right)^n$$

$$e^{-\frac{1}{2}\sum_{k=1}^n x_k^2}dx_1\cdots dx_n = 2^n\int_{\mathbb{R}_+^n} h\left(x_1^2+\cdots+x_n^2;\frac{x_2^2}{\sum_{k=1}^n x_k^2};\cdots;\frac{x_n^2}{\sum_{k=1}^n x_k^2}\right)\left(\frac{1}{\sqrt{2\pi}}\right)^n$$

$$e^{-\frac{1}{2}\sum_{k=1}^n x_k^2}dx_1\cdots dx_n = \int_{\mathbb{R}_+^n}(h\circ g)(x_1;\cdots;x_n)(2\pi)^{-\frac{n}{2}}\exp(-\tfrac{1}{2}g_1(x_1;\cdots;x_n))$$

$$2^n dx_1\cdots dx_n = \int_{\mathbb{R}_+^n}(h\circ g)(x_1;\cdots;x_n)(2\pi)^{-\frac{n}{2}}\exp(-\tfrac{1}{2}g_1(x_1;\cdots;x_n))$$

$$\cdot\frac{(\sum_{k=1}^n x_k^2)^{n-1}}{x_1\cdots x_n}\cdot 2^n\frac{x_1\cdots x_n}{(\sum_{k=1}^n x_k^2)^{n-1}}dx_1\cdots dx_n = \int_{\mathbb{R}_+^n}(h\circ g)(x_1;\cdots;x_n)(2\pi)^{-\frac{n}{2}}$$

$$\exp(-\tfrac{1}{2}g_1(x_1;\cdots;x_n))\cdot\frac{g_1(x_1;\cdots;x_n)^{n-1}\prod_{k=2}^n\frac{1}{\sqrt{g_1(x_1;\cdots;x_n)g_k(x_1;\cdots;x_n)}}}{\sqrt{g_1(x_1;\cdots;x_n)-g_1(x_1;\cdots;x_n)\sum_{k=2}^n g_k(x_1;\cdots;x_n)}}$$

$$\cdot|\det Jg(x_1;\cdots;x_n)|\,dx_1\cdots dx_n = \int_{\mathbb{R}_+\times[0,1]^{n-1}} h(y_1;\cdots;y_n)(2\pi)^{-\frac{n}{2}}$$

$$\exp(-\tfrac{1}{2}y_1)\cdot\frac{y^{n-1}\prod_{k=2}^n\frac{1}{\sqrt{y_1y_k}}}{\sqrt{y_1-y_1\sum_{k=2}^n y_k}}dy_1\cdots dy_n = \int_{\mathbb{R}_+\times[0,1]^{n-1}} h(y_1;\cdots;y_n)(2\pi)^{-\frac{n}{2}}$$

$$y^{\frac{n-1}{2}-\frac{1}{2}}e^{-\frac{1}{2}y_1}\cdot\frac{\prod_{k=2}^n\frac{1}{\sqrt{y_k}}}{\sqrt{1-\sum_{k=2}^n y_k}}dy_1\cdots dy_n$$

where we note

$$\det Jg(g_1;\cdots;g_n) = \det\begin{pmatrix} 2g_1 & 2g_2 & \cdots & 2g_n \\ 2g_1\frac{-g_2^2}{\Sigma^2} & 2g_2\left(\frac{1}{\Sigma}-\frac{g_2^2}{\Sigma^2}\right) & \cdots & 2g_n\frac{-g_2^2}{\Sigma^2} \\ \vdots & & & \vdots \\ 2g_1\frac{-g_n^2}{\Sigma^2} & 2g_2\frac{-g_n^2}{\Sigma^2} & \cdots & 2g_n\left(\frac{1}{\Sigma}-\frac{g_n^2}{\Sigma^2}\right)\end{pmatrix}$$

$$= 2g_1\cdots 2g_n\left(\frac{1}{\Sigma^2}\right)^{n-1}\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ -g_2^2 & \Sigma-g_2^2 & \cdots & -g_2^2 \\ \vdots & & & \\ -g_n^2 & -g_n^2 & \cdots & \Sigma-g_n^2\end{pmatrix}$$

$$= 2^n g_1\cdots g_n\left(\frac{1}{\Sigma^{n-1}}\right)^2\cdot(-g_2^2)\cdots(-g_n^2)\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & -\frac{\Sigma}{g_2^2}+1 & \cdots & 1 \\ \vdots & & & \\ 1 & 1 & \cdots & -\frac{\Sigma}{g_n^2}+1\end{pmatrix}$$

$$= (-1)^{n-1} 2^n g_1 \left(\frac{1}{\Sigma^{n-1}}\right)^2 g_2^3 \cdots g_n^3 \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & -\frac{\Sigma}{g_2^2} & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & -\frac{\Sigma}{g_n^2} \end{pmatrix}$$

$$= (-1)^{n-1} 2^n g_1 \left(\frac{1}{\Sigma^{n-1}}\right)^2 g_2^3 \cdots g_n^3 \cdot 1 \cdot \left(-\frac{\Sigma}{g_2^2}\right) \cdots \left(-\frac{\Sigma}{g_n^2}\right)$$

$$= \left((-1)^{n-1}\right)^2 2^n g_1 \frac{1}{\Sigma^{n-1}} g_2 \cdots g_n = 2^n \frac{g_1 \cdots g_n}{\Sigma^{n-1}}$$

where we use the denotation $\Sigma = \sum_{k=1}^{n} g_k^2$.

Thus the density function for $(X, (\pi_2; \ldots; \pi_n))$ in $\mathbb{R}_+ \times [0,1]^{n-1}$ is

$$(2\pi)^{-\frac{n}{2}} g_1^{\frac{n}{2}-1} e^{-\frac{1}{2}g_1} \frac{1}{\sqrt{1 - \sum_{k=2}^{n} g_k}} \prod_{k=2}^{n} \frac{1}{\sqrt{g_k}} =: p(X, (\pi_2; \ldots; \pi_n))(g_1; \ldots; g_n)$$

Let us now calculate the probability density for X on $\mathbb{R}_+$: let $h: \mathbb{R}_+ \to \mathbb{R}$ be a measurable bounded test function; we have $(P = P_{G_1} \otimes \cdots \otimes P_{G_n})$

$$E_P(h(X)) = \int_{\mathbb{R}^n} h(g_1^2 + \cdots + g_n^2)\, \phi(g_1) \cdots \phi(g_n)\, dg_1 \cdots dg_n = \int_{\mathbb{R}^n} h(g_1^2 + \cdots + g_n^2)\, (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{k=1}^{n} g_k^2} dg_1 \cdots dg_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} (2\pi)^{-\frac{n}{2}} \int_0^\infty h(r^2) e^{-\frac{1}{2}r^2} r^{n-1} dr$$

$$= 2^{1-\frac{n}{2}} \frac{1}{\Gamma(\frac{n}{2})} \int_0^\infty h(s) e^{-\frac{1}{2}s} (s^{\frac{1}{2}})^{n-1} \cdot \frac{1}{2} \frac{1}{\sqrt{s}} ds = \int_0^\infty h(s) \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{1}{2}s} s^{\frac{n}{2}-1} ds$$

where we use n-dimensional spherical coordinates and we make a substitution $s = r^2$; $ds = 2r\,dr$. Hence the probability density function of X in $\mathbb{R}_+$ is $p_X(s) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{1}{2}s} s^{\frac{n}{2}-1}$

Then let us calculate the probability density of $\pi_2; \ldots; \pi_n$ on $[0,1]^{n-1}$. Let $h: [0,1]^{n-1} \to \mathbb{R}$ be a measurable bounded test function; we have $(P = P_{G_1} \otimes \cdots \otimes P_{G_n})$

$$E_P(h(\pi_2; \ldots; \pi_n)) = \int_{\mathbb{R}^n} h\left(\frac{g_2^2}{\sum_{k=1}^{n} g_k^2}; \ldots; \frac{g_n^2}{\sum_{k=1}^{n} g_k^2}\right) (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{k=1}^{n} g_k^2} dg_1 \cdots dg_n = 2^n \int_{\mathbb{R}_+^n} h\left(\frac{g_2^2}{\sum_{k=1}^{n} g_k^2}; \ldots; \frac{g_n^2}{\sum_{k=1}^{n} g_k^2}\right) (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{k=1}^{n} g_k^2} dg_1 \cdots dg_n$$

$$= \int_{\mathbb{R}_+^n} (h \circ f)(g_1; \ldots; g_n)\, (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} f_1(g_1; \ldots; g_n)} \cdot f_1(g_1; \ldots; g_n)^{n-1}$$

$$\frac{1}{\sqrt{g_1(y_1;\cdots;y_n)-g_1(y_1;\cdots;y_n)\sum_{k=2}^{n} g_k(y_1;\cdots;y_n)}}\prod_{k=2}^{n}\frac{1}{\sqrt{g_1(y_1;\cdots;y_n)g_k(y_1;\cdots;y_n)}}$$

$$|\det J(y_1;\cdots;y_n)|\,dy_1\cdots dy_n=\int_{\mathbb{R}_+\times[0,1]^{n-1}} h(y_2;\cdots;y_n)(2\pi)^{-\frac{n}{2}}$$

$$e^{-\frac{1}{2}y_1}\,y_1^{\frac{n-1}{2}-\frac{1}{2}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{1}{\sqrt{y_k}}\,dy_1\cdots dy_n=\int_0^\infty y_1^{\frac{n}{2}-1}e^{-\frac{1}{2}y_1}dy_1\int_{[0,1]^{n-1}}$$

$$h(y_2;\cdots;y_n)(2\pi)^{-\frac{n}{2}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{1}{\sqrt{y_k}}\,dy_2\cdots dy_n=\int_0^\infty (2x)^{\frac{n}{2}-1}$$

$$e^{-x}\cdot 2dx\int_{[0,1]^{n-1}} h(y_2;\cdots;y_n)(2\pi)^{-\frac{n}{2}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{dy_k}{\sqrt{y_k}}$$

$$=2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\int_{[0,1]^{n-1}} h(y_2;\cdots;y_n)(2\pi)^{-\frac{n}{2}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{dy_k}{\sqrt{y_k}}$$

$$=\int_{[0,1]^{n-1}} h(y_2;\cdots;y_n)\frac{\Gamma(\frac{n}{2})}{\pi^{\frac{n}{2}}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{dy_k}{\sqrt{y_k}}$$

where we make the substitution $x=\frac{1}{2}y$, $dx=\frac{1}{2}dy$ and recognize the Γ-function $\Gamma(\frac{n}{2})=\int_0^\infty x^{\frac{n}{2}-1}e^{-x}dx$. Thus the probability density of $(r_2;\dots;r_n)$ on $[0,1]^{n-1}$ is

$$p_{(r_2;\dots;r_n)}(y_2;\cdots;y_n)=\frac{\Gamma(\frac{n}{2})}{\pi^{\frac{n}{2}}}\frac{1}{\sqrt{1-\sum_{k=2}^{n}y_k}}\prod_{k=2}^{n}\frac{1}{\sqrt{y_k}}$$

Hence we get the result

$$p_{(X,(r_2;\dots;r_n))}(y_1;\cdots;y_n)=p_X(y_1)\,p_{(r_2;\dots;r_n)}(y_2;\cdots;y_n)$$

proving that X, $(r_2;\dots;r_n)$ are independent.

Hence the claim is true. $\square$

In the process of the above proof, we also computed the probability density of X on $\mathbb{R}_+$:

$$p_X(s)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}e^{-\frac{1}{2}s}s^{\frac{n}{2}-1};\quad s\in\mathbb{R}_+$$

and the probability density of $(r_2;\dots;r_n)$ on $[0,1]^{n-1}$:

EXERCISES 9.

Option prices and their "greeks"

In mathematical finance, the Black and Scholes model for a stock price $S_t(\omega)$ at time $t \geq 0$ is defined via log-normal distribution (which in other words is the exponential of a Gaussian)

$$S_t(\omega) = S_0 \exp\left(\sigma \sqrt{t}\, G(\omega) - \frac{1}{2}\sigma^2 t\right)$$

where $S_0 \geq 0$ is the stock price at the present time $t=0$ and $G(\omega) \sim N(0,1)$ is a standard Gaussian distribution, with

$$P(G \leq x) = \int_{-\infty}^{x} \phi(y)dy; \quad \phi(y) = \frac{\exp(-\frac{1}{2}y^2)}{\sqrt{2\pi}}$$

An european option with maturity $t>0$ is a random variable defined as a function of the stock price $H(\omega) = h(S_t(\omega))$, where $x \mapsto h(x)$ is measurable.

It follows that the price of the option at the present time $t=0$ is the expectation $c(t; S_0; \sigma) := E_P(h(S_t))$ with respect to the pricing probability P.

1.

Claim: $E_P(S_t) = S_0$ (For this reason, such P is called risk-neutral)

Proof:

Direct proof:

$$E_P(S_t) = \int_{-\infty}^{\infty} S_0 e^{\sigma\sqrt{t}y - \frac{1}{2}\sigma^2 t} \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}} dy = S_0 \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(y-\sigma\sqrt{t})^2}}{\sqrt{2\pi}} dy = S_0 \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx$$

$$= S_0 \cdot 1 = S_0$$

where we make the change of variable $x = y - \sigma\sqrt{t}$, $dx = dy$.

Hence the claim is true. $\square$

2.

Assume that $x \mapsto h(x)$ is differentiable.

Claim: the option price $c(t; S_0; \sigma)$ is differentiable with respect to the parameters $t; S_0$ and σ.

Proof:

Direct proof:

It is clear that the claim is false; a trivial counterexample is given by the function $h(x) = e^x$, for which $c(t; S_0; \sigma) = \infty$ $\forall t, S_0; \sigma \in]0, \infty[$. Thus certain further assumptions of the function h are needed.

So, let us assume the following: (these assumptions justify the change of order in derivation and integration operators by providing appropriate uniform integrability condition for the derivatives in a neighborhood of the parameter vector)

- $x \mapsto h'(e^x) e^{\alpha x} e^{-\beta x^2} \in L^1(\mathbb{R})$ $\forall \alpha \in \mathbb{R}, \forall \beta > 0$
- h' is continuous

This implies $x \mapsto h(e^x) e^{\alpha x} e^{-\beta x^2} \in L^1(\mathbb{R})$ $\forall \alpha \in \mathbb{R}, \forall \beta > 0$ as

$$h(e^x) = \int_0^x h'(e^t) e^t \, dt + h(1)$$

$$\int_{-\infty}^{\infty} |h(e^x) e^{\alpha x} e^{-\beta x^2}| \, dx \le \int_{-\infty}^{\infty} \int_0^x |h'(e^t)| e^t e^{\alpha x} e^{-\beta x^2} \, dt \, dx + |h(1)| \cdot C(\alpha, \beta) \overset{(*)}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty}$$

$$\mathbb{1}(t \ge 0) \, \mathbb{1}(t \le x) |h'(e^t)| e^t e^{\alpha x} e^{-\beta x^2} \, dx \, dt + |h(1)| C(\alpha, \beta) = \int_0^{\infty} |h'(e^t)| e^t \int_t^{\infty} e^{\alpha x}$$

$$e^{-\beta x^2} \, dx \, dt + |h(1)| C(\alpha, \beta) \le \int_0^{\max\{\frac{\alpha}{2\beta}+1; 0\}} |h'(e^t)| e^t \cdot C(\alpha, \beta) \, dt + \int_{\max\{\frac{\alpha}{2\beta}+1, 0\}}^{\infty} |h'(e^t)| e^t$$

$$\int_t^{\infty} (x - \frac{\alpha}{2\beta}) e^{\alpha x} e^{-\beta x^2} \, dx + |h(1)| C(\alpha, \beta) = C(\alpha, \beta) \int_0^{\max\{\frac{\alpha}{2\beta}+1; 0\}} |h'(e^t)| e^t \, dt + \int_{\max\{\frac{\alpha}{2\beta}+1; 0\}}^{\infty} |h'(e^t)| e^t \frac{1}{2\beta} e^{\alpha t} e^{-\beta t^2} d$$

$+ h(1) C(\alpha, \beta) < \infty$ where $C(\alpha, \beta) = \int_{-\infty}^{\infty} e^{\alpha x} e^{-\beta x^2} \, dx < \infty$ (Gaussian integral)

and $\int_0^{\max\{\frac{\alpha}{2\beta}+1; 0\}} |h'(e^t)| e^t \, dt < \infty$ because h' is continuous (and

thus has a maximum in $[1, e^{\max\{\alpha+1; 0\}}]$. Above we use Tonelli in $(*)$ (for positive functions). Also we deduce similarly as in exercise 3 of problem set 8: $\int_{-\infty}^{\infty} g(x) f(x) x e^{-\beta x^2} dx = (0 - 0) + \frac{1}{2\beta} \int_{-\infty}^{\infty} g'(x)$ $f(x) e^{-\beta x^2} dx + \frac{1}{2\beta} \int_{-\infty}^{\infty} g(x) f'(x) e^{-\beta x^2} dx$ (assuming $g(x) f(x) e^{-\beta x^2} \xrightarrow[x \to \pm\infty]{} 0$) and we can apply this formula for $g(x)=e^{\alpha x}$, $f(x)=\mathbb{1}(x>t)$, $f'(x)=\delta_t(x)$ and we get $\int_{-\infty}^{\infty}(x-\frac{\alpha}{2B})e^{\alpha x}f(x)e^{-Bx^2}dx$ $=\frac{1}{2B}\int_{-\infty}^{\infty}e^{\alpha x}\delta_t(x)e^{-Bx^2}dx=\frac{1}{2B}e^{\alpha t}e^{-Bt^2}$. Then we see that

$$\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x-\ln S_0+\frac{1}{2}\sigma^2 t\right)\right)^2\right)=\exp\left(-\frac{1}{2\sigma^2 t}\left(x^2+2x\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}+\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)\right)=\exp\left(-\frac{1}{2\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)\exp\left(-\frac{1}{\sigma^2 t}\ln\left(\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2\sigma^2 t}x^2\right)$$

from which we may deduce that

$$E\left(\left|h'(S_t)\frac{\partial S_t}{\partial t}\right|\right)=\int_{-\infty}^{\infty}|h'(S_0e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t})|\,S_0e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t}\left|\frac{1}{2}\sigma\frac{1}{\sqrt{t}}y-\frac{1}{2}\sigma^2\right|\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}dy=\int_{-\infty}^{\infty}|h'(e^x)|e^x\left|\frac{1}{2t}(x-\ln S_0)-\frac{1}{4}\sigma^2\right|\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x-\ln S_0+\frac{1}{2}\sigma^2 t\right)\right)^2\right)\cdot\frac{1}{\sigma\sqrt{t}}dx\le\frac{1}{\sigma\sqrt{2\pi t}}\frac{1}{2t}\exp\left(-\frac{1}{2\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)\int_{-\infty}^{\infty}|h'(e^x)|e^x\left(|x|+|\ln S_0|+\frac{1}{2}\sigma^2 t\right)\exp\left(-\frac{1}{\sigma^2 t}\ln\left(\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2\sigma^2 t}x^2\right)dx\le\frac{1}{2\sqrt{2\pi}}\frac{1}{\sigma(\alpha)t(\alpha)\sqrt{t(\alpha)}}\exp\left(-\frac{1}{2\sigma(\Omega)^2t(\Omega)}\left(\min\left|\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right|\right)^2\right)\left(\int_0^{\infty}|h'(e^x)|e^x\left(|x|+\max|\ln S_0|+\frac{1}{2}\sigma(\Omega)^2t(\Omega)\right)\exp\left(-\left(\min\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{x^2}{2\sigma(\Omega)^2t(\Omega)}\right)dx+\int_{-\infty}^{0}|h'(e^x)|e^x\left(|x|+\max|\ln S_0|+\frac{1}{2}\sigma(\Omega)^2t(\Omega)\right)\exp\left(-\left(\max\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2\sigma(\Omega)^2t(\Omega)}x^2\right)dx\right)<\infty\quad(|x|\le e^x)$$

$$E\left(\left|h'(S_t)\frac{\partial S_t}{\partial S_0}\right|\right)=\int_{-\infty}^{\infty}|h'(S_0e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t})|\,e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t}\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}dy=\int_{-\infty}^{\infty}|h'(e^x)|e^{x-\ln S_0}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x-\ln S_0+\frac{1}{2}\sigma^2 t\right)\right)^2\right)\frac{1}{\sigma\sqrt{t}}dx$$

$$=\frac{1}{\sigma S_0\sqrt{2\pi t}}\exp\left(-\frac{1}{2\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)\int_{-\infty}^{\infty}|h'(e^x)|e^x\exp\left(\left(-\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2\sigma^2 t}x^2\right)dx\le\frac{1}{\sigma(\alpha)S_0(\alpha)\sqrt{2\pi t(\alpha)}}\exp\left(-\frac{1}{2\sigma(\Omega)^2t(\Omega)}\left(\min\left|\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right|\right)^2\right)\int_{-\infty}^{\infty}|h'(e^x)|e^x\left(\exp\left(-\left(\min\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)+\exp\left(-\left(\max\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\right)\exp\left(-\frac{1}{2\sigma(\Omega)^2t(\Omega)}x^2\right)dx<\infty$$

$$E\left(\left|h'(S_t)\frac{\partial S_t}{\partial\sigma}\right|\right)=\int_{-\infty}^{\infty}|h'(S_0e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t})|\,S_0e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t}\,|\sqrt{t}y-\sigma t|$$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = \int_{-\infty}^{\infty} |h'(e^x)| e^x \left|\frac{1}{\sigma}(x - \ln S_0) - \frac{1}{2}\sigma t\right| \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x - \ln S_0 + \frac{1}{2}\sigma^2 t\right)\right)^2\right) \cdot \frac{1}{\sigma\sqrt{t}}\, dx \le \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2 t}\left(\ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)$$

$$\int_{-\infty}^{\infty} |h'(e^x)| e^x \left(|x| + |\ln S_0| + \frac{1}{2}\sigma^2 t\right) \exp\left(-\frac{1}{\sigma^2 t}\left(\ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right) x\right)$$

$$\exp\left(-\frac{1}{2\sigma^2 t} x^2\right) dx \le \frac{1}{\sigma(\alpha)\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma(\Omega)^2 t(\Omega)}\left(\min\left|\ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right|\right)^2\right)$$

$$\int_{-\infty}^{\infty} |h'(e^x)| e^x \left(|x| + \max|\ln S_0| + \frac{1}{2}\sigma(\Omega)^2 t(\Omega)\right)\left(\exp\left(-\left(\min \frac{1}{\sigma^2 t} \ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right) x\right) + \exp\left(-\left(\max \frac{1}{\sigma^2 t} \ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right) x\right)\right) \exp\left(-\frac{1}{2\sigma(\Omega)^2 t(\Omega)} x^2\right) dx$$

$$< \infty \quad (|x| \le e^x)$$

Thus we see that if $t \in [t(\alpha), t(\Omega)] \subseteq]0, \infty[$, $S_0 \in [S_0(\alpha), S_0(\Omega)] \subseteq]0, \infty[$, $\sigma \in [\sigma(\alpha), \sigma(\Omega)] \subseteq]0, \infty[$, then

$$\sup_{(t; S_0; \sigma) \in [t(\alpha), t(\Omega)] \times [S_0(\alpha), S_0(\Omega)] \times [\sigma(\alpha), \sigma(\Omega)]} \left\{ E\left(\left|h'(S_t) \frac{\partial S_t}{\partial t}\right|\right), E\left(\left|h'(S_t) \frac{\partial S_t}{\partial S_0}\right|\right), E\left(\left|h'(S_t) \frac{\partial S_t}{\partial \sigma}\right|\right) \right\} < \infty$$

Also from the above calculation we can deduce that

$$(1)\; E\left(\left|h'(S_t) \frac{\partial S_t}{\partial ?}\right| \mathbb{1}(\cdot > K)\right) \le Const(?) \int_{-\infty}^{\infty} |h'(e^x)| e^x \left(\exp\left(-\left(\min \frac{1}{\sigma^2 t} \ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right) x\right) + \exp\left(-\left(\max \frac{1}{\sigma^2 t} \ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right) x\right) \right) \exp\left(-\frac{1}{2\sigma(\Omega)^2 t(\Omega)} x^2\right)$$

$$\mathbb{1}(\cdot > K)\, Factor(?, x)\, dx$$

where $? \in \{t; S_0; \sigma\}$ and

$$Const(?) = \exp\left(-\frac{1}{2\sigma(\Omega)^2 t(\Omega)}\left(\min\left|\ln \frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right|\right)^2\right) \begin{cases} \frac{1}{2\sqrt{2\pi}} \frac{t(\alpha)^{-\frac{3}{2}}}{\sigma(\alpha)} ; & ? = t \\ \frac{1}{\sigma(\alpha) S_0(\alpha) \sqrt{2\pi t(\alpha)}} ; & ? = S_0 \\ \frac{1}{\sigma(\alpha)\sqrt{2\pi}} ; & ? = \sigma \end{cases}$$

and

$$Factor(?, x) = \begin{cases} |x| + \max|\ln S_0| + \frac{1}{2}\sigma(\Omega)^2 t(\Omega) ; & ? = t, \sigma \\ 1 ; & ? = S_0 \end{cases}$$

and

$$\mathbb{1}(\cdot > K) = \mathbb{1}(|h'(S_t)\frac{\partial S_t}{\partial ?}| > K)$$

$$\mathbb{1}(: > K) = \mathbb{1}(|h'(e^x)|e^x|Thing(?,x)| > K)$$

where

$$Thing(?,x) = \begin{cases} \frac{1}{2t}(x - \ln S_0) - \frac{1}{4}\sigma^2 ; ? = t \\ \frac{1}{S_0} ; ? = S_0 \\ \frac{1}{\sigma}(x - \ln S_0) - \frac{1}{2}\sigma t ; ? = \sigma \end{cases}$$

Now we want to enlarge the condition in the $\mathbb{1}(: > K)$-factor (because the function being integrated in the right-hand side of (1) is non-negative, this gives us an upper bound). As $|h'(e^x)|e^x \geq 0$, we can do this enlarging by finding an upper bound for for $|Thing(?,x)|$.

$$|Thing(?,x)| \leq Stuff(?,x) := \begin{cases} \frac{1}{2t(\alpha)}(|x| + \max|\ln S_0|) + \frac{1}{4}\sigma(\Omega)^2 ; ? = t \\ \frac{1}{S_0(\alpha)} ; ? = S_0 \\ \frac{1}{\sigma(\alpha)}(|x| + \max|\ln S_0|) + \frac{1}{2}\sigma(\Omega)t(\Omega) ; ? = \sigma \end{cases}$$

So, replacing $\mathbb{1}(: > K)$ from $\mathbb{1}(|h'(e^x)|e^x|Thing(?,x)| > K)$ to $\mathbb{1}(|h'(e^x)|e^x Stuff(?,x) > K)$ gives an upper bound, uniformly in $t \in [t(\alpha), t(\Omega)]$, $S_0 \in [S_0(\alpha), S_0(\Omega)]$, $\sigma \in [\sigma(\alpha), \sigma(\Omega)]$. Looking into the right hand side of (1) with this replacement, we see that (as the right hand side integrand is a $L^1(\mathbb{R})$-function) the right hand side $\to 0$ as $K \to \infty$.

By definition (Gasbarra: Todennäköisyysteorian luennot, kevät 2015 (moniste): Määritelmä 8.0.2) this means that the family

$$A := \{ h'(S_t)\frac{\partial S_t}{\partial ?} \mid t \in [t(\alpha), t(\Omega)], S_0 \in [S_0(\alpha); S_0(\Omega)]; \sigma \in [\sigma(\alpha); \sigma(\Omega)];$$

$$? = t, S_0, \sigma \}$$

is uniformly integrable.

Also, as the map

$? \mapsto \left(h'(S_t)\frac{\partial S_t}{\partial ?}\right)(\omega, ?)$ is continuous for all ω $(? = t; S_0; \sigma)$

and the map

$\omega \mapsto \left(h'(S_t)\frac{\partial S_t}{\partial ?}\right)(\omega, ?)$ is measurable for all $?$ $(? = t; S_0; \sigma)$

we can use Gasbarra: Todennäköisyysteorian luennot, kevät 2015 (moniste): Lause 8.0.6 and say that

$$h(S_t) = \int_{?(\alpha)}^{?} h'(S_t)\frac{\partial S_t}{\partial ?}\, d? + (h(S_t))(? = ?(\alpha)),\quad ? = t; S_0; \sigma$$

has a ?-derivable expectation value $E(h(S_t)) = c(t; S_0; \sigma)$. This holds for all $t \in [t(\alpha), t(\Omega)]$, $S_0 \in [S_0(\alpha), S_0(\Omega)]$; $\sigma \in [\sigma(\alpha), \sigma(\Omega)]$. As for all $t \in]0,\infty[$, $S_0 \in]0,\infty[$, $\sigma \in]0,\infty[$ we can find such intervals $[?(\alpha), ?(\Omega)] \subseteq]0,\infty[$, $? = t; S_0; \sigma$ as above, we conclude that $E(h(S_t)) = c(t; S_0; \sigma)$ is once derivable with respect to any variable $t; S_0; \sigma$ at any point $(t; S_0; \sigma) \in]0,\infty[^3$.

Note that above we know that $E(h(S_t))$ exists as

$$E(h(S_t)) = \int_{-\infty}^{\infty} h\left(S_0 e^{\sigma\sqrt{t}y - \frac{1}{2}\sigma^2 t}\right)\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy = \int_{-\infty}^{\infty} h(e^x)\frac{1}{\sqrt{2\pi}}$$

$$\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x - \ln S_0 + \frac{1}{2}\sigma^2 t\right)\right)^2\right)\frac{1}{\sigma\sqrt{t}}\,dx = \frac{1}{\sigma\sqrt{2\pi t}}\exp\Bigg(-\frac{1}{2\sigma^2 t}$$

$$\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\Bigg)\int_{-\infty}^{\infty} h(e^x)\exp\left(-\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2\sigma^2 t}x^2\right)dx$$

$$< \infty$$

by our previous analysis.

Thus the claim is true. $\square$

3.

Claim: $c(t; S_0; \sigma)$ is differentiable with respect to the parameters $t; S_0$ and σ (also in situations where the function $h(x)$ is not differentiable)

Proof:

Direct proof:

As we saw in the previous exercise 2, the claim is false (with function $h(x) = e^x$ again providing a counterexample) unless suitable conditions are imposed for the function h.

So, let us give the suitable uniform integrability conditions for h.

- $|h(e^x)| e^{\alpha x} e^{-\beta x^2} \in L^1(\mathbb{R}) \ \forall \alpha \in \mathbb{R} \ \forall \beta > 0$.

Let us analyse the situation as follows:

(1) $c(t; S_0; \sigma) = \int_{-\infty}^{\infty} h(\exp(\log S_0 - \frac{1}{2}\sigma^2 t + \sigma\sqrt{t}\, y))\, \phi(y) dy = \frac{1}{\sigma\sqrt{t}}$
$\int_{-\infty}^{\infty} h(\exp(x))\, \phi(\frac{1}{\sigma\sqrt{t}}(x - \log S_0 + \frac{1}{2}\sigma^2 t)) dx$

where we use the change of variables

$x = \log S_0 - \frac{1}{2}\sigma^2 t + \sigma\sqrt{t}\, y$

$\Rightarrow y = \frac{1}{\sigma\sqrt{t}}(x - \log S_0 + \frac{1}{2}\sigma^2 t)$

From this we get

$\frac{\partial}{\partial ?}(h(e^x)\phi(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))) = h(e^x)(-(\frac{1}{\sigma\sqrt{t}}(x + \log(\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})))$
$\phi(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})) \cdot \frac{\partial}{\partial ?}(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})) = -\frac{1}{\sigma^2 t} h(e^x)$
$(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})\, \phi(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\, \frac{\partial}{\partial ?}\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}$

for $? = t; S_0; \sigma$ and above we note that the density of the standard Gaussian distribution satisfies $\frac{d}{dx}\phi(x) = -x\phi(x)$.

We note that we can add $\phi(x)$:

$h(e^x)\phi(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})) = h(e^x)\phi(\frac{1}{\sigma\sqrt{t}}(x + \log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\, \phi^{-1}(x)\phi(x)$

so as to say that

$$c(t;S_0;\sigma)=E[h(e^G)\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\phi(G)^{-1}]$$

We also get

$$\frac{\partial}{\partial ?}(h(e^G)\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\phi(G)^{-1})=-\frac{1}{\sigma^2 t}h(e^G)\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))$$
$$\phi(G)^{-1}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})\frac{\partial}{\partial ?}\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}=:Y(t;S_0;\sigma;\omega)$$

Thus, as

$? \mapsto Y(t;S_0;\sigma;\omega)$ is continuous $\forall\omega$, $?=t;S_0;\sigma$ (for $(t;S_0;\sigma)\in]0,\infty[^3$)

$\omega\mapsto Y(t;S_0;\sigma;\omega)$ is measurable $\forall t;S_0;\sigma\in]0,\infty[$

we only need to show that the families

$$\{-\frac{1}{\sigma^2 t}h(e^G)\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\phi(G)^{-1}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})\frac{\partial}{\partial ?}(\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})$$
$$?\in[?(\alpha),?(\Omega)]\}=:F(?)$$

where $?=t;S_0;\sigma$ and $[?(\alpha),?(\Omega)]\subseteq]0,\infty[$, $?=t;S_0;\sigma$ are uniformly integrable to apply Gasbarra: Todennäköisyysteorian luennot, kevät 2015 (moniste): Lause 8.0.6 which tells us that

$$c(t;S_0;\sigma)=E[\int_{?(\alpha)}^{?}\frac{\partial}{\partial ?}(h(e^G)\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\phi(G)^{-1})\,d?+(h(e^G)$$
$$\phi(\frac{1}{\sigma\sqrt{t}}(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}))\phi(G)^{-1})(?(\alpha);\hat{?};\hat{\hat{?}})]$$

is differentiable at $?\in]?(\alpha),?(B)[$, where $?=t;S_0;\sigma$ and $\{?,\hat{?},\hat{\hat{?}}\}=\{t;S_0;\sigma\}$. From this it follows that $c(t;S_0;\sigma)$ is differentiable with respect to $t;S_0;\sigma$ at all points $t;S_0;\sigma\in]0,\infty[$.

We see that, as $-\frac{1}{\sigma^2 t}\frac{\partial}{\partial ?}(\log\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0})$ is just a bounded constant factor (when $?=t;S_0;\sigma$ and $t\in[t(\alpha),t(\Omega)]$, $S_0\in[S_0(\alpha),S_0(\Omega)]$, $\sigma\in[\sigma(\alpha),\sigma(\Omega)]$) it does not affect the uniform integrability of the families $F(?)$, so we can ignore it. This also means that we only need to show uniform integrability for

$$h(e^G)\phi\left(\frac{1}{\sigma\sqrt{t}}\left(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)\right)\phi(G)^{-1}\left(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)$$

to cover all cases.

We note that

$$\left|h(e^G)\phi\left(\frac{1}{\sigma\sqrt{t}}\left(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)\right)\phi(G)^{-1}\left(G+\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)\right| \le |h(e^G)|$$
$$\exp\left(-\frac{1}{2}\frac{1}{\sigma^2 t}\left(\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)^2\right)\exp\left(-\frac{1}{\sigma^2 t}\left(\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)G\right)\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\right.\right.$$
$$\left.\left.+1\right)G^2\right)\left(|G|+\left|\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right|\right) \le \exp\left(-\frac{1}{2}\frac{1}{\sigma(\Omega)^2 t(\Omega)}\left(\min\left|\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right|\right)^2\right)$$
$$|h(e^G)|\left(\exp\left(-\frac{1}{\sigma(\Omega)^2 t(\Omega)}\left(\min\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)G\right)+\exp\left(-\frac{1}{\sigma(\alpha)^2 t(\alpha)}\right.\right.$$
$$\left.\left.\left(\max\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)G\right)\right)\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma(\Omega)^2 t(\Omega)}+1\right)G^2\right)\left(|G|+\max\left|\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right|\right)$$

where the bound is uniform over the family and since the bound is L^1-integrable, the family is L^1-integrable and uniformly integrable (Gasbarra: Todennäköisyysteorian luennot, kevät 2015 (moniste): Lemma 8.0.3).

The bound is L^1-integrable as

$$\int_{-\infty}^{\infty}|h(e^x)|\left(\exp\left(-\frac{1}{\sigma(\Omega)^2 t(\Omega)}\left(\min\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)x\right)+\exp\left(-\frac{1}{\sigma(\alpha)^2 t(\alpha)}\right.\right.$$
$$\left.\left.\left(\max\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right)x\right)\right)\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma(\Omega)^2 t(\Omega)}+1\right)x^2\right)\left(|x|+\max\left|\log\frac{e^{\frac{1}{2}\sigma^2 t}}{s_0}\right|\right)$$
$$\frac{e^{-x^2}}{\sqrt{2\pi}}\,dx<\infty$$

by assumption.

Hence, as said above, the claim is true. $\square$

5.

Let us compute the option price $c(t;s_0;\sigma)$ and the sensitivity parameters

$$\frac{\partial c(t;s_0;\sigma)}{\partial t},\quad \frac{\partial c(t;s_0;\sigma)}{\partial s_0},\quad \frac{\partial c(t;s_0;\sigma)}{\partial \sigma}$$

(these derivatives are referred in the math-finance literature as "greeks") in the following two cases

a) $h(S_t)=(S_t-K)^+$, which is an european call-option with strike price $K>0$,

b) $h(S_t)=(K-S_t)^+=(S_t-K)^-=K-S_t-(S_t-K)^+$ which is an european put-option with strike price $K>0$.

a) We have

$$c(t;S_0;\sigma)=\int_{-\infty}^{\infty}\left(S_0 e^{\sigma\sqrt{t}y-\frac{1}{2}\sigma^2 t}-K\right)^+\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy=\int_{-\infty}^{\infty}(e^x-K)^+$$

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\left(x-\ln S_0+\frac{1}{2}\sigma^2 t\right)\right)^2\right)\frac{1}{\sigma\sqrt{t}}dx=\frac{1}{\sigma\sqrt{2\pi t}}\exp\Big(-\frac{1}{2}$$

$$\frac{1}{\sigma^2 t}\Big(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)^2\Big)\int_{\ln K}^{\infty}(e^x-K)\exp\left(-\frac{1}{\sigma^2 t}\ln\left(\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)\exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2 t}\right)dx$$

Now, let us introduce the identity

$$\int_t^{\infty}e^{\alpha x}e^{-\beta x^2}dx=e^{\frac{\alpha^2}{4\beta}}\int_t^{\infty}e^{-\beta(x-\frac{\alpha}{2\beta})^2}dx=e^{\frac{\alpha^2}{4\beta}}\int_{\sqrt{2\beta}(t-\frac{\alpha}{2\beta})}^{\infty}e^{-\frac{1}{2}y^2}\frac{1}{\sqrt{2\beta}}dy$$

$$=\frac{\sqrt{2\pi}}{\sqrt{2\beta}}e^{\frac{\alpha^2}{4\beta}}\left(1-\Phi\left(\sqrt{2\beta}\left(t-\frac{\alpha}{2\beta}\right)\right)\right)$$

where Φ is the cumulative distribution function of the standard normal distribution. Applying this above we get

$$c(t;S_0;\sigma)=\frac{1}{\sigma\sqrt{2t\left(\frac{1}{2}\frac{1}{\sigma^2 t}\right)}}\exp\left(-\frac{1}{2}\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)\Big(\exp\Big(\frac{1}{4\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}$$

$$\Big(1-\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)^2\Big)\Big(1-\Phi\Big(\sqrt{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big(\ln K-\frac{1}{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big(1-\frac{1}{\sigma^2 t}$$

$$\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)\Big)\Big)\Big)-K\exp\Big(\frac{1}{4\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big(-\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)^2\Big)\Big(1-\Phi\Big(\sqrt{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}$$

$$\Big(\ln K-\frac{1}{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big(-\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)\Big)\Big)\Big)\Big)=\exp\Big(\frac{1}{2}\sigma^2 t-\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)$$

$$\Big(1-\Phi\Big(\frac{1}{\sigma\sqrt{t}}\Big(\ln K-\sigma^2 t+\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)\Big)\Big)-K\Big(1-\Phi\Big(\frac{1}{\sigma\sqrt{t}}\Big(\ln K+\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)\Big)\Big)$$

$$=S_0-K+\left(K\Phi\left(\frac{\ln K}{\sigma\sqrt{t}}+\frac{1}{2}\sigma\sqrt{t}-\frac{1}{\sigma\sqrt{t}}\ln S_0\right)-S_0\Phi\left(\frac{\ln K}{\sigma\sqrt{t}}-\frac{1}{2}\sigma\sqrt{t}-\frac{\ln S_0}{\sigma\sqrt{t}}\right)\right)$$

$$=S_0-K+\left(K\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sigma\sqrt{t}\right)-S_0\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)\right)$$

The sensitivity parameters are

$$\frac{\partial c(t;S_0;\sigma)}{\partial t} = K\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sigma\sqrt{t}\right)^2}\left(-\frac{1}{2}\frac{1}{\sigma}t^{-\frac{3}{2}}\ln\frac{K}{S_0}+\frac{1}{4}\sigma\frac{1}{\sqrt{t}}\right)$$

$$-S_0\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)^2}\left(-\frac{1}{2}\frac{1}{\sigma}t^{-\frac{3}{2}}\ln\frac{K}{S_0}-\frac{1}{4}\sigma\frac{1}{\sqrt{t}}\right)=\frac{1}{\sqrt{2\pi}}$$

$$\Big(\frac{1}{2}\frac{1}{\sigma}t^{-\frac{3}{2}}\ln\frac{S_0}{K}\,e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}\left(Ke^{-\frac{1}{2}\ln\frac{K}{S_0}}-S_0e^{+\frac{1}{2}\ln\frac{K}{S_0}}\right)$$

$$+\frac{1}{4}\sigma\frac{1}{\sqrt{t}}e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}\left(Ke^{-\frac{1}{2}\ln\frac{K}{S_0}}+S_0e^{+\frac{1}{2}\ln\frac{K}{S_0}}\right)\Big)=\frac{1}{2\sigma\sqrt{2\pi t}}$$

$$e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}\left(\frac{1}{t}\ln\frac{S_0}{K}\left(K\frac{\sqrt{S_0}}{\sqrt{K}}-S_0\frac{\sqrt{K}}{\sqrt{S_0}}\right)+\frac{1}{2}\sigma^2\left(K\frac{\sqrt{S_0}}{\sqrt{K}}+S_0\frac{\sqrt{K}}{\sqrt{S_0}}\right)\right)$$

$$=\frac{1}{2}\sigma\sqrt{\frac{KS_0}{2\pi t}}\,e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}$$

$$\frac{\partial c(t;S_0;\sigma)}{\partial S_0}=1+K\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sigma\sqrt{t}\right)^2}\left(\frac{1}{\sigma\sqrt{t}}\cdot\left(-\frac{1}{S_0}\right)\right)-1\cdot\Phi\Big($$

$$\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\Big)-S_0\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)^2}\cdot\left(\frac{1}{\sigma\sqrt{t}}\cdot\left(-\frac{1}{S_0}\right)\right)$$

$$=1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)+\frac{1}{\sigma\sqrt{2\pi t}}e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}$$

$$\left(e^{+\frac{1}{2}\ln\frac{K}{S_0}}-\frac{K}{S_0}e^{-\frac{1}{2}\ln\frac{K}{S_0}}\right)=1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)$$

$$\frac{\partial c(t;S_0;\sigma)}{\partial\sigma}=K\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sigma\sqrt{t}\right)^2}\left(-\frac{1}{\sigma^2\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sqrt{t}\right)$$

$$-S_0\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{K}{S_0}-\frac{1}{2}\sigma\sqrt{t}\right)^2}\left(-\frac{1}{\sigma^2 t}\ln\frac{K}{S_0}-\frac{1}{2}\sqrt{t}\right)=\frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{1}{2}\Big(\frac{1}{\sigma^2 t}$$

$$\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\Big)\Big)\left(\frac{1}{\sigma^2\sqrt{t}}\ln\frac{S_0}{K}\left(Ke^{-\frac{1}{2}\ln\frac{K}{S_0}}-S_0e^{+\frac{1}{2}\ln\frac{K}{S_0}}\right)+\frac{1}{2}\sqrt{t}\Big(Ke^{-\frac{1}{2}\ln\frac{K}{S_0}}\right.$$

$$\left.+S_0e^{+\frac{1}{2}\ln\frac{K}{S_0}}\Big)\right)=\sqrt{\frac{tKS_0}{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)\right)$$

b) We have (almost similarly as before)

$$c(t;S_0;\sigma)=\frac{1}{\sigma\sqrt{2\pi t}}e^{-\frac{1}{2}\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2}\int_{-\infty}^{\ln K}(K-e^x)\exp\left(-\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\right)$$

$$\exp\left(-\frac{1}{2}\frac{1}{\sigma^2 t}x^2\right)dx=\frac{1}{\sigma\sqrt{2\pi t}}e^{-\frac{1}{2}\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2}\int_{-\ln K}^{\infty}(K-e^{-x})\exp\Big(\frac{1}{\sigma^2 t}$$

$$\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)x\Big)\exp\left(-\frac{1}{2}\frac{1}{\sigma^2 t}x^2\right)dx=\frac{1}{\sigma\sqrt{2t\cdot\left(\frac{1}{2}\frac{1}{\sigma^2 t}\right)}}\exp\left(-\frac{1}{2}\frac{1}{\sigma^2 t}\left(\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)^2\right)$$

$$\Big(K\exp\Big(\frac{1}{4\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big)\Big(\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\Big)^2\Big)\Big(1-\Phi\Big(\sqrt{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\Big(-\ln K-\frac{1}{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}$$

$\left(\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right))))-\exp\left(\frac{1}{4\cdot\frac{1}{2}\sigma^2 t}\left(\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}-1\right)^2\right)\left(1-\Phi\left(\sqrt{2\cdot\frac{1}{2}\ldots}\right.\right.$

$\left(-\ln K-\frac{1}{2\cdot\frac{1}{2}\frac{1}{\sigma^2 t}}\left(\frac{1}{\sigma^2 t}\ln\left(\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)-1\right)\right)))\Big)^+ = K\left(1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\left(-\ln K\right.\right.\right.$

$\left.-\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}\right)))-\exp\left(\frac{1}{2}\sigma^2 t\left(-2\frac{1}{\sigma^2 t}\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}+1\right)\right)\left(1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\left(-\ln\right.\right.\right.$

$\left.-\ln\frac{e^{\frac{1}{2}\sigma^2 t}}{S_0}+\sigma^2 t\right)))=K\left(1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{S_0}{K}-\frac{1}{2}\sigma\sqrt{t}\right)\right)-S_0\left(1-\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\right.\right.$

$\left.+\frac{1}{2}\sigma\sqrt{t}\right))=K-S_0+\left(S_0\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{S_0}{K}+\frac{1}{2}\sigma\sqrt{t}\right)-K\Phi\left(\frac{1}{\sigma\sqrt{t}}\ln\frac{S_0}{K}-\frac{1}{2}\sigma\sqrt{t}\right.\right.$

From this we get that

$$c(t;S_0;\sigma)_{b)}=-c(t;S_0;-\sigma)_{a)}$$

where a) and b) refer to the c-functions of parts a) and b), respectively. Thus we get the sensitivity parameters

$$\frac{\partial c(t;S_0;\sigma)_{b)}}{\partial t}=-\frac{\partial c(t;S_0;-\sigma)_{a)}}{\partial t}=\frac{1}{2}\sigma\sqrt{\frac{KS_0}{2\pi t}}\,e^{-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma^2 t\right)}$$

$$\frac{\partial c(t;S_0;\sigma)_{b)}}{\partial S_0}=-\frac{\partial c(t;S_0;-\sigma)_{a)}}{\partial S_0}=\Phi\left(\frac{1}{-\sigma\sqrt{t}}\ln\frac{K}{S_0}+\frac{1}{2}\sigma\sqrt{t}\right)-1$$

$$\frac{\partial c(t;S_0;\sigma)_{b)}}{\partial \sigma}=-\frac{\partial c(t;S_0;-\sigma)}{\partial \sigma}\cdot(-1)=\sqrt{\frac{tKS_0}{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2 t}\left(\ln\frac{K}{S_0}\right)^2+\frac{1}{4}\sigma\ldots\right.\right.$$

The european call-option $(S_t(\omega)-K)^+$ gives to the option holder the right (without obligation) to buy from his counterpart at the maturity time t one share at the pre-determined price K. If at the time of maturity t the market price of the stock $S_t(\omega)$ is higher than the stri[ke] price K, the option holder will exercise her right, buy[ing] one share at price K and selling in the market at mark[et] price $S_t(\omega)$, making a profit of $(S_t(\omega)-K)^+$. If at maturity time t $S_t(\omega)\le K$, the call-option becomes worthless, and the option holder does not make profit.

Analogously the european put-option $(S_t(\omega)-K)^-$ gives to the option holder the right (without obligation) to sell to the counterpart one share at the pre-determined price K. Th[e] put-call parity between european put and call options is the equ[ation]

$$S_t(\omega)-K=(S_t(\omega)-K)^+-(K-S_t(\omega))^+$$

Since the contract on the left side at time $t=0$ has price S_0-K, we have the following parity relation between the put and call option prices at time $t=0$:

$$S_0-K=c((S_t-K)^+)-c((K-S_t)^+)$$

For simplicity we have assumed that a riskless investment has zero interest rate, equivalently all values are discounted and expressed in present-time values.

EXERCISES 10.

1.

Let $H \subseteq L^2(\Omega, \mathcal{F}, P)$ be a closed subspace. We follow the definition of Gasbarra: Lecture notes in probability theory fall semester 2015 (moodle): Proposition 3.11

a) Claim: the L^2-projection π_H is a linear operator: when $X, Y \in L^2(P)$, $a, b \in \mathbb{R}$

$$\pi_H(aX+bY) = a\pi_H(X) + b\pi_H(Y).$$

Proof:

Direct proof:

From $E[(X-\pi_H(X))W]=0 \ \forall W \in H$ (similarly for Y) we get

$$(1)\begin{cases} E[((aX+bY)-\pi_H(aX+bY))W]=0 \\ E[((aX-a\pi_H(X))+(bY-b\pi_H(Y)))W]=0 \end{cases}$$

for all $W \in H$ (note that

$E[((aX-a\pi_H(X))+(bY-b\pi_H(Y)))W] = a\,E[(X-\pi_H(X))W] + b\,E[(Y-\pi_H(Y))W] = a\cdot 0 + b\cdot 0 = 0$)

By taking the difference of above equations (1) we get

$$E[(a\pi_H(X)+b\pi_H(Y)-\pi_H(aX+bY))W]=0$$

for all $W \in H$. But choosing $W = a\pi_H(X)+b\pi_H(Y)-\pi_H(aX+bY) \in H$ we get

$$E[(a\pi_H(X)+b\pi_H(Y)-\pi_H(aX+bY))^2]=0$$

$\Rightarrow a\pi_H(X)+b\pi_H(Y)-\pi_H(aX+bY)=0$ P-almost surely

$\Rightarrow a\pi_H(X)+b\pi_H(Y)=\pi_H(aX+bY)$ (projection defined only P-almost surely)

Hence the claim is true. $\square$

b) Claim: the L^2 projection is idempotent: $(\pi_H)^2 = \pi_H$; meaning that when $Y \in H$, $\pi_H Y = Y$

Proof:

Direct proof:

Because for any $X \in L^2(P)$, $\pi_H(X) \in H$, it is enough to show $\pi_H(Y) = Y$ $\forall Y \in H$ to deduce the idempotence $\pi_H^2 = \pi_H$.

So, let $Y \in H$. Then $Y \in H$ is an element of H such that

1. $E[(Y-Y)^2] = 0 = \Delta^2 = \inf_{W \in H} E[(Y-W)^2]$ (as $E[(Y-W)^2] \geq 0$ $\forall Y, W \in H$)
2. $E[(Y-Y)W] = 0$ $\forall W \in H$

Thus Y satisfies the definition of the L^2-projection: $\pi_H Y = Y$. (See Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Proposition 9.1.1). Note that projection is defined only up to P-null events.
Hence the claim is true. □

c) Claim: the projection does not increase the L^2 norm:

$$\|X\|_{L^2(P)} \geq \|\pi_H(X)\|_{L^2(P)}$$

Proof:

Direct proof:

Let $X \in L^2(P)$. Then

$$\|X\|^2_{L^2(P)} = E[X^2] = E[((X - \pi_H(X)) + \pi_H(X))^2] = E[(X - \pi_H(X))^2]$$

$$+ 2E[(X - \pi_H(X))\pi_H(X)] + E[\pi_H(X)^2] = E[(X - \pi_H(X))^2] + 2 \cdot 0$$

$$+ E[\pi_H(X)^2] \geq 0 + E[\pi_H(X)^2] = \|\pi_H(X)\|^2_{L^2(P)}$$

$$\Rightarrow \|X\|_{L^2(P)} \geq \|\pi_H(X)\|_{L^2(P)}$$

Hence the claim is true. □

In general these properties characterize orthogonal (general projection operators lack property c)) projection operators.

(*) We will use in the multivariate case the following extension of the linear predictor formula from example 9.1.1 in the lecture notes (Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste)):

When $X=(X_1;\dots;X_T)^T \in L^2(\Omega,\mathcal{F},P)$ then the following multivariate formula holds: for $Y=(Y_1;\dots;Y_d)^T$ another random variable in $L^2(P)$, in matrix vector notations

$$\hat{Y}=E[Y]+(X-E[X])^T(E[XX^T]-E[X]E[X]^T)^{-1}(E[XY^T]$$

$$-E[X]E[Y]^T)=E[Y]+(X-E[X])^T \operatorname{Cov}(X,X)^{-1}\operatorname{Cov}(X,Y)$$

where M^{-1} denotes the inverse of a matrix M and $\operatorname{Cov}(X,Y)_{i,j} = E[X_iY_j]-E[X_i]E[Y_j]$ is the covariance between X_i and Y_j and $\hat{Y}$ is the $L^2(P)$-projection of Y to the linear span of $\{1;X_1;\dots;X_T\}$.

2.

Let $G\sim N(0,1)$ be a standard Gaussian variable with probability density $\phi(g)=(2\pi)^{-\frac{1}{2}}\exp(-\frac{1}{2}g^2)$, and let $x\mapsto f(x)$ be a differentiable function with $E[f(G)^2]<\infty$ and $E[|f'(G)|]<\infty$.

a) Claim: $\hat{f}(G)=E[f(G)]+E[f'(G)]G$ is the best affine approximation of $f(G)$ based on G in least square sense, meaning that $\hat{a}=E[f(G)]$ and $\hat{b}=E[f'(G)]$ minimize the mean square error $E[(f(G)-(a+bG))^2]$.

Proof:

Direct proof:

By definition, the (mean square error minimizing) best affine approximation is given by the L^2-projection to the subspace span$\{1,G\}$. By Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 this projection is given by (note $E[G]=0$)

$$\hat{f}(G) = E[f(G)] + \frac{E[f(G)G]}{E[G^2]} G = E[f(G)] + E[f'(G)]G$$

where we use Gaussian integration by parts formula (Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Proposition 7.0.6) and we note $E[G^2]=1$ for standard Gaussian.

Hence the claim is true. □

b) Now we consider the same linear approximation in the multivariate case. Let $G(\omega) = (G_1(\omega); \dots; G_T(\omega))^T \in \mathbb{R}^T$, where the coordinates $G_t(\omega)$ are independent and identically distributed standard Gaussian random variables. Let $f: \mathbb{R}^T \to \mathbb{R}$ be differentiable with $E[f(G_1; \dots; G_n)^2] < \infty$ and

$$E\left[\left|\frac{\partial}{\partial x_t} f(G_1; \dots; G_T)\right|\right] < \infty.$$

Claim:

$$\hat{f}(G_1; \dots; G_T) = E[f(G_1; \dots; G_T)] + \sum_{t=1}^{T} E\left[\frac{\partial}{\partial x_t} f(G_1; \dots; G_T)\right] G_t$$

is the best affine approximation of $f(G_1; \dots; G_T)$ in the linear span of $\{1; G_1; \dots; G_T\}$ with coefficients minimizing the mean square error

$$E\left[\left(f(G_1; \dots; G_T) - \left(c_0 + \sum_{t=1}^{T} c_t G_t\right)\right)^2\right]$$

Proof:

Direct proof:

Again as in part a), the best approximation is given by L^2-projection to span$\{1, G_1; \dots; G_T\}$. By the multivariate extension of Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 ((*) above) the L^2-projection is given by

$$\hat{f}(G_1; \dots; G_T) = E[f(G_1; \dots; G_T)] + (G - E[G])^T \mathrm{Cov}(G, G)^{-1} \mathrm{Cov}(G, f(G))$$

We have

$E[G] = (E[G_1]; \dots; E[G_T])^T = 0$

$Cov(G,G)_{i,j} = E[G_i G_j] - E[G_i]E[G_j] = \begin{cases} 0\cdot 0 - 0\cdot 0, i \neq j = \delta_{i,j} \\ 1 - 0\cdot 0, i = j \end{cases}$

$\Rightarrow Cov(G,G) = \mathbb{1} = Cov(G,G)^{-1}$

$Cov(G, f(G))_{i,1} = E[G_i f(G_1;\dots;G_T)] - E[G_i]E[f(G_1;\dots;G_T)]$

$= E[\frac{\partial}{\partial x_i} f(G_1;\dots;G_T)] - 0\cdot E[f(G_1;\dots;G_T] = E[\frac{\partial}{\partial x_i} f(G_1;\dots;G_T)]$

Putting all this together we get

$f(G_1;\dots;G_T) = E[f(G_1;\dots;G_T)] + \sum_{\ell=1}^{T} E[\frac{\partial}{\partial x_\ell} f(G_1;\dots;G_T)] G_\ell$

Hence the claim is true. $\square$

Above we use the multidimensional Gaussian integration by parts formula:

$E[G_i f(G_1;\dots;G_T] = \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{T-1} (\sqrt{2\pi})^{-(T-1)} \exp(-\frac{1}{2}\sum_{\substack{j=1\\ j\neq i}}^{T} x_j^2) (\int_{-\infty}^{\infty} x_i$

$f(x_1;\dots;x_T) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2} dx_i) dx_1 \cdots \widehat{dx_i} \cdots dx_T \overset{(*)}{=} \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{T-1} (\sqrt{2\pi})^{-(T-1)}$

$\exp(-\frac{1}{2}\sum_{\substack{j=1\\ j\neq i}}^{T} x_j^2) (\int_{-\infty}^{\infty} (\frac{\partial}{\partial x_\ell} f(x_1;\dots;x_T)) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2} dx_i) dx_1 \cdots \widehat{dx_\ell} \cdots dx_T$

$= E[\frac{\partial}{\partial x_\ell} f(G_1;\dots;G_T)]$

where we note $f \in L^2$; $\omega \mapsto G_i(\omega) \in L^2$ imply $E[|G_i f(G_1;\dots;G_T)|] < \infty$ by Hölder's inequality and thus the expectation is a iterated integral by Fubini's theorem (note: $f \in L^2 \Rightarrow$ boundary terms vanish in integration by parts). Note that when applying the integration by parts formula in $(*)$ above, we need the condition $E[|\frac{\partial}{\partial x_i} f(G_1;\dots;G_T)|] < \infty$, because we need the boundary integral $\int_{-\infty}^{\infty} \frac{d}{dx_i}(e^{-\frac{1}{2}x_i^2} f(x_1;\dots;x_T) dx_i$ to be such that

- $e^{-\frac{1}{2}x_i^2} f(x_1;\dots;x_T)$ admits derivative everywhere (obviously OK)

- Derivative of $e^{-\frac{1}{2}x_\ell^2} f(x_1; \dots; x_T)$ is integrable (follows from $f \in L^2(\phi(x)dx)$, $\frac{\partial}{\partial x_\ell} f \in L(\phi(x)dx)$)

Then we may apply fundamental theorem of calculus and get that

$$\int_{-\infty}^{\infty} \frac{d}{dx_\ell} (e^{-\frac{1}{2}x_\ell^2} f(x_1; \dots; x_T)) dx_\ell = \lim_{x_\ell \to \infty} (e^{-\frac{1}{2}x_\ell^2} f(x_1; \dots; x_T))$$

$$- \lim_{x_\ell \to -\infty} (e^{-\frac{1}{2}x_\ell^2} f(x_1; \dots; x_T)) = 0$$

as $f \in L^2(\phi(x)dx)$.

c) Next we consider the correlated case: let $A = (A_{s,t})$ be a nonsingular $T \times T$ matrix, $G = (G_1; \dots; G_T)$ with independent, identically distributed standard Gaussian coordinates as before and let $X = (X_1; \dots; X_T) = AG$ with coordinates

$$X_s = \sum_{t=1}^{T} A_{s,t} G_t$$

We have seen that the random vector X is Gaussian with zero mean and covariance matrix $\Sigma = AA^T$. Let $f(x_1; \dots; x_T)$ be a differentiable function with

$$E[f(X_1; \dots; X_T)^2] < \infty$$

and

$$E[|\frac{\partial}{\partial x_\ell} f(X_1; \dots; X_T)|] < \infty.$$

Let us compute the coefficients of the best linear approximation $\hat{f}(X_1; \dots; X_T)$ of $f(X_1; \dots; X_T)$ in the linear span of $\{1; X_1; \dots; X_T\}$ minimizing the mean square error

$$E[(f(X_1; \dots; X_T) - (c_0 + \sum_{\ell=1}^{T} c_\ell X_\ell))^2].$$

Again as in parts a), b) the best approximation is given by the L^2-projection to span $\{1; X_1; \dots, X_T\}$. By the multivariate extension of Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 ((*) above) the L^2-projection is given by

$$\hat{f}(X) = E[f(X)] + (X - E[X])^T Cov(X, X)^{-1} Cov(X, f(X))$$

The coefficients are

$$E[g(X_1;\dots;X_T)],$$

$$E[X]=0$$

$$Cov(X,X)^{-1}=\Sigma^{-1}=(A^{-1})^T A^{-1}$$

$$Cov(X,g(X))_{t,1}=E[X_t g(X_1;\dots;X_T)]-E[X_t]E[g(X_1;\dots;X_T)]$$

$$=\sum_{s=1}^{T}A_{t,s}E[G_s g(\sum_{r=1}^{T}A_{1,r}G_r;\dots;\sum_{r=1}^{T}A_{T,r}G_r)]-0\cdot E[g(X_1;\dots;X_T)]$$

$$=\sum_{s=1}^{T}A_{t,s}E[\frac{\partial}{\partial G_s}g(\sum_{r=1}^{T}A_{1,r}G_r;\dots;\sum_{r=1}^{T}A_{T,r}G_r)]=\sum_{s=1}^{T}A_{t,s}\sum_{r=1}^{T}A_{r,s}E[$$

$$\frac{\partial}{\partial x_r}g(\sum_{p=1}^{T}A_{1,p}G_p;\dots;\sum_{p=1}^{T}A_{T,p}G_p)]=\sum_{r=1}^{T}(AA^T)_{t,r}E[\frac{\partial}{\partial x_r}g(X_1;\dots;X_T)]$$

Hence we have

$$\hat{g}(X)=E[g(X)]+X^T\Sigma^{-1}\Sigma E[\nabla g(X)]=E[g(X)]+X^TE[\nabla g(X)]$$

3.

Let $N(\omega)$ be a Poisson(λ) distributed random variable with parameter $\lambda>0$, where

$$P_\lambda(N=k)=e^{-\lambda}\frac{\lambda^k}{k!}\quad\text{for } k\in\mathbb{N}_0$$

and $(g(k):k\in\mathbb{N}_0)$ be a sequence with $E[g(N)^2]<\infty$.

a) Claim: $\hat{g}(N)=E[g(N)]+E[g(N+1)-g(N)](N-\lambda)$ is the best estimator of $g(N)$ depending on N, with coefficients minimizing the mean square error

$$E[(g(N)-(a+bN))^2]$$

Proof:

Direct proof:

The **best estimator** is given by L^2-projection to span$\{1,G\}$. By Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 it is given by

$$f(N) = E[f(N)] + \frac{Cov(N f(N))}{Var(N)}(N - E[N])$$

We know that $E[N] = \lambda$, $E[N^2] = \lambda^2 + \lambda$ and the Stein equation for Poisson-λ random variable:

$$E[N f(N)] = \lambda E[f(N+1)]$$

Hence

$$Cov(N f(N)) = E[N f(N)] - E[N]E[f(N)] = \lambda E[f(N+1)] - \lambda E[f(N)]$$

$$Var(N) = E[N^2] - E[N]E[N] = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Hence

$$f(N) = E[f(N)] + (E[f(N+1)] - E[f(N)])(N - \lambda)$$

Thus the claim is true. $\square$

b) Let now $N = (N_1, \ldots, N_T) \in \mathbb{N}_0^T$ where the coordinates N_t are independent and Poisson(λ_t) distributed for $t = 1, \ldots, T$, respectively, with $\lambda_t > 0$ (possibly different).

Let $f: \mathbb{N}_0^T \to [0, \infty[$ be a function with $E[f(N_1, \ldots, N_T)^2] < \infty$.

Claim: $f(N_1, \ldots, N_T) = E[f(N_1, \ldots, N_T)] + \sum_{t=1}^{T} E[(f(N_1, \ldots, N_{t-1}, N_t + 1, N_{t+1}, \ldots, N_T)$

$- f(N_1, \ldots, N_{t-1}, N_t, N_{t+1}, \ldots, N_T)](N_t - \lambda_t)$ is the best affine

approximation of $f(N_1, \ldots, N_T)$ in the linear span of $\{1, N_1, \ldots, N_T\}$ with coefficients minimizing the mean square error

$$E[(f(N_1, \ldots, N_T) - (c_0 + \sum_{t=1}^{T} c_t N_t))^2]$$

Proof:

Direct proof:

The best approximation is given by the L^2-projection to span$\{1, N_1, \ldots, N_T\}$. By the multivariate extension of Gasbarra: Lecture notes in probability theory, fall semester 2015 (monisite):

Example 9.1.1 ((*) above) the projection is given by

$$\hat{f}(N) = E[f(N)] + (N - E[N])^T \operatorname{Cov}(N,N)^{-1} \operatorname{Cov}(N, f(N))$$

We have

$$E[N] = (E[N_1]; \ldots; E[N_T])^T = (\lambda_1; \ldots, \lambda_T)^T$$

$$\operatorname{Cov}(N,N)_{s,t} = E[N_s N_t] - E[N_s]E[N_t] = (\lambda_s^2 + \lambda_s - \lambda_s^2)\delta_{s,t} = \lambda_s \delta_{s,t}$$

$$\operatorname{Cov}(N, f(N))_{t,1} = E[N_t f(N)] - E[N_t]E[f(N)] = \lambda_t E[f(N_1; \ldots; N_{t-1}; N_t + 1; N_{t+1}; \ldots; N_T)] - \lambda_t E[f(N)]$$

where we use the Stein equation (see a)-part). Note $f \in L^2$, $N_t \in L^2 \Rightarrow E[N_t f(N)]$ can be calculated as iterated sums (Hölder + Fubini), so the one-dimensional Stein equation of part a) can be easily generalized to multiple dimensions.

Hence

$$\hat{f}(N_1; \ldots; N_T) = E[f(N_1; \ldots; N_T)] + \sum_{t=1}^{T} E[f(N_1; \ldots; N_{t-1}; N_t + 1; N_{t+1}; \ldots; N_T) - f(N_1; \ldots; N_{t-1}; N_t; N_{t+1}; \ldots; N_T)](N_t - \lambda_t)$$

Hence the claim is true. □

5.

Let G be a standard Gaussian random variable.

For f differentiable with derivative satisfying $E[|\partial f(G)|] < \infty$, we define the adjoint operator $f \mapsto \partial^* f$ with $(\partial^* f)(x) = x f(x) - \partial f(x)$.

a) Let us use the Gaussian integration by parts formula (Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Proposition 7.0.6) together with the product rule of calculus

$$\partial(fh) = f\partial h + h \partial f$$

to prove the following extension of the Gaussian integration by parts formula:

Claim: when $E[f(G)^2]<\infty$ and $E[(\partial f(G))^2]<\infty$, and for another differentiable h with $E[(\partial h(G))^2]<\infty$,

$$E[h(G)\partial f(G)] = E[f(G)\partial^* h(G)]$$

Proof:

Direct proof:

$$E[h(G)\partial f(G)] = E[\partial(hf)(G)] - E[f(G)\partial h(G)] = E[G h(G) f(G)]$$

$$-E[f(G)\partial h(G)] = E[f(G)\partial^* h(G)]$$

Here we demand additional assumption $E[h(G)^2]<\infty$ to make sure the integrals $E[h(G)\partial f(G)]$, $E[f(G)\partial h(G)]$ (Hölder) and thus $E[\partial(hf)(G)]$ exist.

Hence the claim is true. $\square$

∂^* is the adjoint of the derivative operator ∂ in the space $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \phi(x)dx)$ where the integration measure is the standard Gaussian distribution on $\mathbb{R}$.

b) We define the (unnormalized) Hermite polynomials as $h_0(x)=1$, and by induction $h_n(x) = ((\partial^*)^n(1))(x) = \partial^* h_{n-1}(x)$.

Let us compute the first six Hermite polynomials $h_n(x)$ for $n=1,2,3,4,5$.

$$h_0(x) = 1$$

$$h_1(x) = x\cdot 1 - \partial_x(1) = x$$

$$h_2(x) = x\cdot x - \partial_x x = x^2-1$$

$$h_3(x) = x(x^2-1) - \partial_x(x^2-1) = x^3-3x$$

$$h_4(x) = x(x^3-3x) - \partial_x(x^3-3x) = x^4-6x^2+3$$

$$h_5(x) = x(x^4-6x^2+3) - \partial_x(x^4-6x^2+3) = x^5-10x^3+15x$$

c) Claim: $E[h_n(G)]=0$

Proof:

Direct proof:

The claim is actually false as $E[h_0(G)]=E[1]=1\neq 0$.

We have, by part a) ($h_n\in L^2$ and $\partial h_n\in L^2$ as they are polynomials; $x\mapsto 1\in L^2$, $\partial(x\mapsto 1)\in L^2$):

$$E[h_n(G)]=E[\partial^* h_{n-1}(G)]=E[1\cdot\partial^* h_{n-1}(G)]=E[h_{n-1}(G)\cdot(\partial(1))(G)]$$

$$=E[h_{n-1}(G)\cdot 0]=0 \quad \forall n\geq 1$$

Hence the claim is true. □

d) Claim: $E[h_m(G)h_n(G)]=m!\delta_{m,n}$ and in particular the random variables $h_m(G)$ and $h_n(G)$ are orthogonal in $L^2(\Omega,\mathcal{F},P)$.

Proof:

Direct proof:

Let $m,n\in\mathbb{N}_0$, $m\leq n$. Then ($h_m, \partial^i h_m\in L^2\ \forall m,i\in\mathbb{N}_0$ as they are polynomials)

$$E[h_m(G)h_n(G)]=E[h_m(G)\,\partial^* h_{n-1}(G)]=E[\partial h_m(G)\,h_{n-1}(G)]$$

$$=E[\partial h_m(G)\,\partial^* h_{n-2}(G)]=\dots=E[\partial^n h_m(G)\,h_0(G)]=E[\partial^n h_m(G)\cdot 1]$$

$$=E[\partial^n h_m(G)]$$

Now, as is clear from the construction of h_m (b)-part), h_m is a polynomial of degree m whose leading term is x^m. Thus

$$\partial^n h_m(x)=\begin{cases}0, & n>m\\ m!, & n=m\end{cases}$$

So $E[h_m(G)h_n(G)]=E[m!\delta_{m,n}]=m!\delta_{m,n}$. Hence the claim is true. □

5.

Let f be a function with n derivatives $\partial^n f$, such that $\partial^k f(G) \in L^2(\Omega, \mathcal{F}, P)$ for $k = 0, 1, 2, \ldots, n$.

Claim: $\hat{f}(G) = E[f(G)] + \sum_{k=1}^{n} \frac{E[\partial^k f(G)]}{k!} h_k(G)$ is the best polynomial

approximation of $f(G)$ in the linear span of $\{h_0(G) = 1; h_1(G) = G; \ldots; h_n(G)\}$ with coefficients minimizing the mean square error

$E[(f(G) - (\sum_{k=0}^{n} c_k h_k(G)))^2]$.

Proof:

Direct proof:

The best approximation is given by the L^2-projection to span$\{h_0(G); h_1(G); \ldots; h_n(G)\}$, which by the multivariate extension of Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 ((*) above) is given by

$\hat{f}(G) = E[f(G)] + (h(G) - E[h(G)])^T Cov(h(G), h(G))^{-1} Cov(h(G), f(G))$,

where we denote $h(G) = (h_1(G); \ldots; h_n(G))^T$ (Note: $h_0(G) = 1$). We have

$E[h(G)] = 0$ (exercise 4 part c))

$Cov(h(G), h(G))_{m,n} = E[h_m(G) h_n(G)] - E[h_m(G)] E[h_n(G)]$

$= m! \delta_{m,n} - 0 \cdot 0 = m! \delta_{m,n}$ (exercise 4 part c), d))

$Cov(h(G), f(G))_{m,1} = E[h_m(G) f(G)] - E[h_m(G)] E[f(G)] = E[f(G)$

$\partial^* h_{m-1}(G)] - 0 \cdot E[f(G)] = E[\partial f(G) \partial^* h_{m-2}(G)] = \ldots = E[\partial^m f(G)$

$h_0(G)] = E[\partial^m f(G)]$

Hence

$\hat{f}(G) = E[f(G)] + \sum_{k=1}^{n} \frac{E[\partial^k f(G)]}{k!} h_k(G)$

Hence the claim is true. $\square$

Similar polynomial approximations can be computed in the multivariate case, and also for Poisson random variables, in that case using some polynomials other than Hermite polynomials, and also in the combined case where the linear span contains the polynomials of both Gaussian and Poisson random variables.

6.

We compute linear projections with Bernoulli random variables. Let X be a binary random variable with

$$P(X=1)=1-P(X=0)=p$$

and $p\in[0,1]$.

a) Claim: the best affine approximation of $f(X)$ for $f:\{0,1\}\to\mathbb{R}$ in the linear span of $\{1,X\}$ in the mean square sense is given by

$$\hat{f}(X)=E[f(X)]+(f(1)-f(0))(X-p)$$

Proof:

Direct proof:

The best approximation is given by the L^2-projection, which by Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 is given by

$$\hat{f}(X)=E[f(X)]+\frac{Cov(f(X),X)}{Var(X)}(X-E[X])$$

We have

$$E[X]=p=E[X^2]$$

$$Var(X)=E[X^2]-E[X]E[X]=p-p^2$$

$$Cov(f(X),X)=E[Xf(X)]-E[X]E[f(X)]=p\cdot 1\cdot f(1)-p(p\cdot f(1)+(1-p)\cdot f(0))$$

$$=(p-p^2)f(1)-(p-p^2)f(0)=(p-p^2)(f(1)-f(0))$$

Hence

$$\tilde{f}(X) = E[f(X)] + (f(1) - f(0))(X - p)$$

Hence, the claim is true. $\square$

b) Claim: $\tilde{f}(X) = f(X)$ (so actually in this case the approximation is exact)

Proof:

Direct proof:

$$\tilde{f}(X) = E[f(X)] + (f(1) - f(0))(X - p) = (pf(1) + (1-p)f(0)) + (f(1) - f(0))(X - p)$$

$$= f(0) + (f(1) - f(0))X = \begin{cases} f(1), & X = 1 \\ f(0), & X = 0 \end{cases} = f(X)$$

Hence the claim is true. $\square$

c) Let $X_1, \dots, X_T$ independent random variables with

$$P(X_t = 1) = 1 - P(X_t = 0) = p_t$$

and $p_t \in [0,1]$, and $f : \{0,1\}^T \to \mathbb{R}$.

Claim: best affine approximation of $f(X)$ in the linear span of $\{1; X_1; \dots; X_T\}$ in the mean square sense is given by

$$\tilde{f}(X_1; \dots; X_T) = E[f(X)] + \sum_{t=1}^{T} E[f(X_1; \dots; X_{t-1}; 1; X_{t+1}; \dots; X_T) - f(X_1; \dots;$$
$$X_{t-1}; 0; X_{t+1}; \dots; X_T)](X_t - p_t)$$

Proof:

Direct proof:

The best approximation is given by L^2-projection to span$\{1; X_1; \dots; X_T\}$; by the multivariate extension of Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Example 9.1.1 ((*) above) it is given by

$\tilde{g}(X) = E[g(X)] + (X - E[X])^T \mathrm{Cov}(X,X)^{-1}\,\mathrm{Cov}(X, g(X))$ ..

where we denote $X = (X_1; \dots; X_T)$. We have

$E[X] = (E[X_1]; \dots; E[X_T])^T = (p_1; \dots; p_T)^T$

$\mathrm{Cov}(X,X)_{s,t} = E[X_s X_t] - E[X_s]E[X_t] = \begin{cases} 0, & s \neq t \\ p_s - p_s \cdot p_s = p_s - p_s^2; & s = t \end{cases}$

$\mathrm{Cov}(X, g(X))_{t,1} = E[X_t g(X)] - E[X_t]E[g(X)] = \sum_{w \in \{0,1\}^T} w_t\, g(w)$
$p_1^{w_1}(1-p_1)^{1-w_1} \cdots p_T^{w_T}(1-p_T)^{1-w_T} - p_t \sum_{w \in \{0,1\}^T} g(w)\, p_1^{w_1}(1-p_1)^{1-w_1} \cdots$
$\cdot\, p_T^{w_T}(1-p_T)^{1-w_T} = \sum_{(w_1;\dots;w_{t-1};w_{t+1};\dots;w_T) \in \{0,1\}^{T-1}} p_1^{w_1}(1-p_1)^{1-w_1} \cdots p_{t-1}^{w_{t-1}}$
$(1-p_{t-1})^{1-w_{t-1}}\, p_{t+1}^{w_{t+1}}(1-p_{t+1})^{1-w_{t+1}} \cdot \ldots \cdot p_T^{w_T}(1-p_T)^{1-w_T} \big(0 \cdot g(w_1;\dots;w_{t-1};$
$0; w_{t+1}; \dots; w_T) \cdot p_t^0 (1-p_t)^{1-0} + 1 \cdot g(w_1; \dots; w_{t-1}; 1; w_{t+1}; \dots; w_T)\, p_t^1 (1-p_t)^{1-1}$
$- p_t\, g(w_1; \dots; w_{t-1}; 0; w_{t+1}; \dots; w_T)\, p_t^0 (1-p_t)^{1-0} - p_t\, g(w_1; \dots; w_{t-1};$
$1; w_{t+1}; \dots; w_T)\, p_t^1 (1-p_t)^{1-1} \big) = \sum_{(w_1;\dots;w_{t-1};w_{t+1};\dots;w_T) \in \{0,1\}^{T-1}}$
$p_1^{w_1}(1-p_1)^{1-w_1} \cdot \ldots \cdot p_{t-1}^{w_{t-1}}(1-p_{t-1})^{1-w_{t-1}} \cdot p_{t+1}^{w_{t+1}}(1-p_{t+1})^{1-w_{t+1}} \cdot \ldots \cdot p_T^{w_T}$
$(1-p_T)^{1-w_T}(p_t - p_t^2)\big(g(w_1; \dots; w_{t-1}; 1; w_{t+1}; \dots; w_T) - g(w_1; \dots; w_{t-1};$
$0; w_{t+1}; \dots; w_T)\big) = (p_t - p_t^2) \sum_{(w_1;\dots;w_{t-1};w_{t+1};\dots;w_T) \in \{0,1\}^{T-1}}$
$\big(g(w_1; \dots; w_{t-1}; 1; w_{t+1}; \dots; w_T) - g(w_1; \dots; w_{t-1}; 0; w_{t+1}; \dots; w_T)\big)$
$p_1^{w_1}(1-p_1)^{1-w_1} \cdot \ldots \cdot p_{t-1}^{w_{t-1}}(1-p_{t-1})^{1-w_{t-1}} \cdot p_{t+1}^{w_{t+1}}(1-p_{t+1})^{1-w_{t+1}} \cdots$
$\cdot\, p_T^{w_T}(1-p_T)^{1-w_T} \overset{!}{=} (p_t - p_t^2)\, E[g(X_1; \dots; X_{t-1}; 1; X_{t+1}; \dots; X_T)$
$- g(X_1; \dots; X_{t-1}; 0; X_{t+1}; \dots; X_T)]$

Hence

$\tilde{g}(X) = E[g(X)] + \sum_{t=1}^{T} E[g(X_1; \dots; X_{t-1}; 1; X_{t+1}; \dots; X_T) - g(X_1; \dots; X_{t-1}; 0;$

$X_{t+1}, \dots, X_T)] (X_t - \eta_t)$

Hence the claim is true. $\square$

7.

Claim: the space $L^\infty(\Omega, \mathcal{F}, P)$ equipped with the essential supremum norm is complete.

$\|X\|_\infty = P\text{-esssup}_{\omega \in \Omega} |X(\omega)| = \inf\{K \in \mathbb{R} \mid P(|X| > K) = 0\}$

Proof:

Direct proof:

We note that, for $\varepsilon > 0$

$\|X\|_\infty + \|Y\|_\infty \geq (K - \varepsilon) + (K' - \varepsilon) \geq \|X + Y\|_\infty - 2\varepsilon$

where $K, K' \in \mathbb{R}$ are such that $P(|X| > K) = 0 = P(|Y| > K')$, $K - \varepsilon \leq \|X\|_\infty$, $K' - \varepsilon \leq \|Y\|_\infty$.

As the above holds for $\forall \varepsilon > 0$, we have $\|X\|_\infty + \|Y\|_\infty \geq \|X + Y\|_\infty$.

Hence it is clear that $L^\infty(\Omega, \mathcal{F}, P)$ is a vector space (obviously, $\|aX\|_\infty = a\|X\|_\infty$), and given that we understand that $X \in L^\infty(\Omega, \mathcal{F}, P)$ is defined only up to P-null deviation, it is clear that $L^\infty(\Omega, \mathcal{F}, P)$ is a normed space with norm $\|\cdot\|_\infty$. Thus $L^\infty(\Omega, \mathcal{F}, P)$ naturally is a metric space.

Let us show completeness. Let $(X_n)_{n \in \mathbb{N}} \subseteq L^\infty(\Omega, \mathcal{F}, P)$ be a Cauchy sequence in $\|\cdot\|_\infty$ norm. This is equivalent to the following:

(☆) $\forall \varepsilon > 0 \ \exists N_\varepsilon \in \mathbb{N} \ \forall n, m \geq N_\varepsilon : |X_n(\omega) - X_m(\omega)| < \varepsilon$ P-almost surely.

Now, noting that countable union and intersection of $P=1$ sets is $P=1$ set, we can see that $(X_n(\omega))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ is P-almost surely a Cauchy sequence, for the set of $\omega \in \Omega$ for which $(X_n(\omega))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ is a Cauchy sequence is

$$\bigcap_{k\in\mathbb{N}} \bigcap_{n\geq N_{\frac{1}{k}}} \bigcap_{m\geq N_{\frac{1}{k}}} \{\omega\in\Omega \mid |X_n(\omega)-X_m(\omega)|<\tfrac{1}{k}\}$$

where $N_{\frac{1}{k}}$ is the smallest of the numbers N_ε for $\varepsilon=\frac{1}{k}$ satisfying the condition (*) above.

So, as $\mathbb{R}$ is complete and P-almost surely $(X_n(\omega))_{n\in\mathbb{N}}\subseteq\mathbb{R}$ is a Cauchy-sequence, P-almost surely exists $X_\infty(\omega)=\lim_{n\to\infty}X_n(\omega)$, $X_\infty(\omega)\in\mathbb{R}$. As X_∞ is the (P-almost sure) pointwise limit on $(X_n)_{n\in\mathbb{N}}\subseteq L^\infty(\Omega,\mathcal{F},P)$, X_∞ is measurable.

Now, when $\varepsilon>0$, let $N_\varepsilon\in\mathbb{N}$ be as above; for $n\geq N_\varepsilon$ we have

$$\|X_n-X_\infty\|_\infty \leq \underbrace{\|X_n-X_{N_\varepsilon}\|}_{|X_n(\omega)-X_{N_\varepsilon}(\omega)|<\varepsilon \text{ a.s.}}+\underbrace{\|X_{N_\varepsilon}-X_\infty\|_\infty}_{|X_{N_\varepsilon}(\omega)-\lim_{n\to\infty}X_n(\omega)|<\varepsilon \text{ a.s.}} < \varepsilon+\varepsilon=2\varepsilon$$

Thus $\|X_n-X_\infty\|_\infty \xrightarrow[n\to\infty]{} 0$; as $(X_n)_{n\in\mathbb{N}}\subseteq L^\infty(\Omega,\mathcal{F},P)$ is Cauchy,

this means that $\|X\|_\infty<\infty$ so $X_\infty\in L^\infty(\Omega,\mathcal{F},P)$ and also $X_n\xrightarrow[n\to\infty]{} X_\infty$ in $L^\infty(\Omega,\mathcal{F},P)$.

Hence. $L^\infty(\Omega,\mathcal{F},P)$ is complete.

Hence the claim is true. $\square$

EXERCISES 11

In all problems the random variables live in the probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subseteq \mathcal{F}$ is a sub-σ-algebra of $\mathcal{F}$.

1.

Let $X \in L^1(\Omega, \mathcal{F}, P)$.

Claim: $|E[X|\mathcal{G}](\omega)| \leq E[|X||\mathcal{G}](\omega)$ (P-almost surely)

Proof:

Direct proof:

By Jensen's inequality, applied for convex function $g: \mathbb{R} \to \mathbb{R}$, $g(x) = |x|$:

$$g(E[X|\mathcal{G}]) = |E[X|\mathcal{G}]| \leq E[g(X)|\mathcal{G}] = E[|X||\mathcal{G}]$$

From which the claim follows as the conditional expectation is defined P-almost surely. (For $X \in L^1(\Omega, \mathcal{F}, P)$, the conditional expectation is well-defined).

See Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Section 10.3 part 6.

Hence the claim is true. $\square$

2.

Let $X, Y \in L^2(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra. The covariance between the random variables X and Y was defined as

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

and we define now the conditional covariance between the random variables X and Y given the information in the σ-algebra $\mathcal{G}$ as the random variable

$$Cov(X,Y|\mathcal{G}) = E[XY|\mathcal{G}] - E[X|\mathcal{G}]E[Y|\mathcal{G}]$$

Claim: $Cov(X,Y) = E[Cov(X,Y|\mathcal{G})] + Cov(E[X|\mathcal{G}], E[Y|\mathcal{G}])$

Proof:

Direct proof:

Let us use the following property of the conditional expectation (Gasbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Section 10.3, part 2) $E[E[X|\mathcal{G}]] = E[X]$:

$$E[Cov(X,Y|\mathcal{G})] + Cov(E[X|\mathcal{G}], E[Y|\mathcal{G}]) = E[E[XY|\mathcal{G}]]$$

$$-E[E[X|\mathcal{G}]E[Y|\mathcal{G}]] + E[E[X|\mathcal{G}]E[Y|\mathcal{G}]] - E[E[X|\mathcal{G}]]E[E[Y|\mathcal{G}]]$$

$$= E[XY] - E[X]E[Y] = Cov(X,Y).$$

Hence the claim is true. $\square$

3.

On a probability space $(\Omega, \mathcal{F}, P)$, let $A \in \mathcal{F}$ with $P(A) > 0$. For a random variable $X(\omega) \geq 0$ P-almost surely, we have defined the (elementary) conditional expectation of X conditioned on the event A as the deterministic constant

$$E[X|A] = \frac{E[X \mathbb{1}_A]}{P[A]}.$$

Consider now the sub-σ-algebra $\mathcal{G} = \sigma(A) = \{\emptyset, \Omega, A, A^c\} \subseteq \mathcal{F}$.

a) Let us compute the conditional expectation of X conditioned on the σ-algebra $\mathcal{G}$ according to the Kolmogorov definition of $E[X|\mathcal{G}]$.

Let us understand that $E[X|\mathcal{G}]$ is well-defined even if $E[X] = \infty$ $(X \geq 0)$.

We note that for a $\mathcal{G}$-measurable $Y \geq 0$ we can write

$$Y = a\mathbb{1}_A + b\mathbb{1}_{A^c}$$

where $a, b \in [0, \infty]$. We have

$$E[Y\mathbb{1}_A] = aP[A], \quad E[Y\mathbb{1}_{A^c}] = bP[A^c], \quad E[Y\mathbb{1}_\Omega] = aP[A] + bP[A^c]$$

and of course $E[Y \mathbb{1}_\emptyset]=0$. For X we have similarly

$$E[X\mathbb{1}_A]=E[X|A]P[A], \quad E[X\mathbb{1}_{A^c}]=E[X|A^c]P[A^c], \quad E[X\mathbb{1}_\Omega]=E[X|A]P[A]+E[X|A^c]P[A^c]$$

and $E[X\mathbb{1}_\emptyset]=0$. So, choosing

$$a=E[X|A]$$

$$b=E[X|A^c]$$

we see that

$$Y=E[X|A]\mathbb{1}_A+E[X|A^c]\mathbb{1}_{A^c}$$

satisfies the Kolmogorov definition of conditional expectation $E[X|\mathcal{G}]$, also

$$Y=E[X|A]\mathbb{1}_A+E[X|A^c]\mathbb{1}_{A^c}=E[X|\mathcal{G}]$$ (P-almost surely as conditional expectation is defined only P-almost surely).

b) Let X be a random variable with cumulative distribution function $F_X(t)=P(X\leq t)$, and let $Y=\mathbb{1}_{]a,b]}\circ X$. ($a\in[-\infty,\infty[$, $b\in]-\infty,\infty[$)

Let us compute the conditional expectation $E[X|\mathcal{G}]$ with respect to the σ-algebra $\mathcal{G}=\sigma(Y)$.

We have

$$\mathcal{G}=\sigma(Y)=\{\emptyset,\Omega,A,A^c\}$$

where

$$A=\{\omega\in\Omega \mid Y(\omega)=1\}=\{\omega\in\Omega \mid X(\omega)\in]a,b]\}.$$

By a)-part,

$$E[X|\mathcal{G}]=E[X|A]\mathbb{1}_A+E[X|A^c]\mathbb{1}_{A^c}=\frac{E[X\mathbb{1}_A]}{P[A]}\mathbb{1}_A+\frac{E[X\mathbb{1}_{A^c}]}{P[A^c]}\mathbb{1}_{A^c}$$

$$= \frac{E[X \mathbb{1}_{X \in]a,b]}]}{F(b)-F(a)} \mathbb{1}_{X \in]a,b]} + \frac{E[X \mathbb{1}_{X \notin]a,b]}]}{1-F(b)+F(a)} \mathbb{1}_{X \notin]a,b]}$$

c) Let $\tau \geq 0$ be a non-negative random variable with λ-exponential distribution such that $P[\tau > s] = e^{-\lambda s}$ for $s \geq 0$. We can interpret τ as a random time. For $t \geq 0$ let us compute the conditional expectation $E[\tau | \sigma(\tau \wedge t)]$ where $(\tau \wedge t)(\omega) = \min\{\tau(\omega), t\}$.

Applying part b) with $X = \tau, a = -\infty, b = t, Y = \mathbb{1}_{]-\infty, t]} \circ \tau$ we get

$$E[\tau | \sigma(\tau \wedge t)] = \frac{E[\tau \mathbb{1}_{\tau \leq t}]}{F(t) - 0} \mathbb{1}_{\tau \leq t} + \frac{E[\tau \mathbb{1}_{\tau > t}]}{1 - F(t) + 0} \mathbb{1}_{\tau > t}$$

$$= \int_0^t \lambda e^{-\lambda x} dx \cdot \frac{\mathbb{1}_{\tau \leq t}}{1 - e^{-\lambda t}} + \int_t^\infty \lambda e^{-\lambda x} dx \frac{\mathbb{1}_{\tau > t}}{1 - (1 - e^{-\lambda t})} = (1 - e^{-\lambda t}) \frac{\mathbb{1}_{\tau \leq t}}{1 - e^{-\lambda t}}$$

$$+ (e^{-\lambda t} - 0) \frac{\mathbb{1}_{\tau > t}}{e^{-\lambda t}} = \mathbb{1}_{\tau \leq t} + \mathbb{1}_{\tau > t} = 1$$

where we agree $\frac{0}{0} = 1$ (in case $t = 0$ and we condition on an event of 0 probability, which we should not do).

4.

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, and $S_n = X_1 + \cdots + X_n$.

Let us compute $E[X_1 | \sigma(S_n)]$.

We note that by linearity of conditional expectation,

$$\sum_{k=1}^n E[X_k | \sigma(S_n)] = E[\sum_{k=1}^n X_k | \sigma(S_n)] = E[S_n | \sigma(S_n)] = S_n$$

where we also note $S_n \in \sigma(S_n)$ (see Garbaczewski: Lecture notes in probability theory, fall semester 2015 (moniste): Section 10.3, part 4; $Y = S_n$, $X = 1$).

By symmetry we see that $E[X_k | \sigma(S_n)] = E[X_1 | \sigma(S_n)]$ $\forall k = 1, \ldots, n$ (For $j, k \in \{1, \ldots, n\}$, $E[E[X_j | \sigma(S_n)] \mathbb{1}_A] = E[X_j \mathbb{1}_A] = E[X_k \mathbb{1}_A] = E[E[X_k | \sigma(S_n)] \mathbb{1}_A]$, also $E[X_j | \sigma(S_n)] = E[X_k | \sigma(S_n)]$).

Thus we get $E[X_1 | \sigma(S_n)] = \frac{1}{n} S_n$.

5.

Let $X, Y \in L^2(\Omega, \mathcal{F}, P)$.

Claim: Cauchy-Schwarz inequality for the conditional expectation:

$$E[XY|\mathcal{G}]^2 \leq E[X^2|\mathcal{G}]E[Y^2|\mathcal{G}]$$

Proof:

Direct proof:

Let $X_n = n \wedge |X|$; $Y_n = n \wedge |Y|$; $n \in \mathbb{N}$ ($\wedge = \min$). Then $0 \leq X_n \leq n$, $0 \leq Y_n \leq n$ $\forall n \in \mathbb{N}$ so that $X_n Y_n \in L^1(\Omega, \mathcal{F}, P)$ (and thus the conditional expectations $E[X_n Y_n|\mathcal{G}]$, $E[X_n^2|\mathcal{G}]$, $E[Y_n^2|\mathcal{G}]$ exist $\forall n \in \mathbb{N}$)

Let $t \in \mathbb{R}$. We have

$$0 \leq (tX_n + Y_n)^2 = t^2 X_n^2 + 2t X_n Y_n + Y_n^2$$

By taking conditional expectations (linearity, positivity)

$$(1)\quad 0 \leq t^2 E[X_n^2|\mathcal{G}] + 2t E[X_n Y_n|\mathcal{G}] + E[Y_n^2|\mathcal{G}]$$

Now, pointwise, the quadratic equation

$$t^2 E[X_n^2|\mathcal{G}](\omega) + 2t E[X_n Y_n|\mathcal{G}](\omega) + E[Y_n^2|\mathcal{G}](\omega) = 0$$

must have at most one solution (since (1) does not attain negative values), so its discriminant must be non-positive:

$$(2E[X_n Y_n|\mathcal{G}](\omega))^2 - 4E[X_n^2|\mathcal{G}](\omega)E[Y_n^2|\mathcal{G}](\omega) \leq 0$$

Now, as $0 \leq X_n^2(\omega) \uparrow X^2(\omega)$; $0 \leq Y_n^2(\omega) \uparrow Y^2(\omega)$; $0 \leq X_n Y_n(\omega) \uparrow |X(\omega)Y(\omega)|$ $\forall \omega \in \Omega$, by the monotone convergence theorem (Yesbarra: Lecture notes in probability theory, fall semester 2015 (moniste): Section 10.3, part 1)

$$E[|XY||\mathcal{G}]^2 = \lim_{n\to\infty} E[X_n Y_n|\mathcal{G}]^2 \leq \lim_{n\to\infty} E[X_n^2|\mathcal{G}]E[Y_n^2|\mathcal{G}]$$

$$= E[X^2|\mathcal{G}]E[Y^2|\mathcal{G}]$$

where we note that $X, Y \in L^2(\Omega, \mathcal{F}, P) \Rightarrow |XY|, X^2, Y^2 \in L^1(\Omega, \mathcal{F}, P)$ so the conditional expectations $E[|XY|/\mathcal{G}], E[X^2|\mathcal{G}], E[Y^2|\mathcal{G}]$ exist.

Now, by exercise 1,

$$E[XY/\mathcal{G}]^2 = |E[XY|\mathcal{G}]|^2 \leq E[|XY|/\mathcal{G}]^2 \leq E[X^2|\mathcal{G}]E[Y^2|\mathcal{G}].$$

Hence the claim is true. $\square$

Note: all conditional expectations above should be understood as defined P-almost surely.

6.

Claim: conditional Chebyshev inequality: when $X \in L^2(\Omega, \mathcal{F}, P)$ and Y is $\mathcal{G}$-measurable, with $P(Y>0)=1$,

$$P(|X|>Y|\mathcal{G}) \leq \frac{E[X^2|\mathcal{G}]}{Y^2} \quad P\text{-almost surely}$$

Proof:

Direct proof:

Let $X \in L^2(\Omega, \mathcal{F}, P)$ and $Y \in \mathcal{G}$, $P(Y>0)=1$. Then

$$Y^2 \mathbb{1}_{|X|>Y} \leq X^2$$

$$\Rightarrow Y^2 E[\mathbb{1}_{|X|>Y}|\mathcal{G}] = Y^2 P(|X|>Y|\mathcal{G}) \leq E[X^2|\mathcal{G}]$$

$$\Rightarrow P(|X|>Y|\mathcal{G}) \leq \frac{E[X^2|\mathcal{G}]}{Y^2} \quad P\text{-almost surely } (Y>0 \ P\text{-almost surely})$$

where we take conditional expectations and use conditional expectation positivity and then we take out the known $Y^2 \in \mathcal{G}$ from the conditional expectation (see Gasbarra. Lecture notes in probability theory, fall semester 2015 (monisite): Section 10.3, part 5).

Note: $X^2 \in L^1(\Omega, \mathcal{F}, P) \Rightarrow E[X^2|\mathcal{G}]$ exists. We also assume $Y^2 \in L^1(\Omega, \mathcal{G}, P)$.

Hence the claim is true. $\square$

Note: In particular the above claim holds when $Y(\omega) \equiv y > 0$ is a deterministic constant.

Note: Above we assume $P(Y>0)=1$ so we do not divide by 0, and we assume $Y^2 \in L^1(\Omega, \mathcal{G}, P)$ so that $E[Y^2 \mathbb{1}_{|X|>Y} | \mathcal{G}]$ exists. Both are necessary assumptions, if we demand that conditional expectation be integrable; if we settle for only measurability of conditional expectation, we can drop $Y^2 \in L^1(\Omega, \mathcal{G}, P)$ (as $Y^2 \geq 0$).

Lemma

Let $G(\omega) \sim N(0,1)$ standard Gaussian, with probability density

$$\phi(x) = \frac{\exp(-\frac{x^2}{2})}{\sqrt{2\pi}}, \quad x \in \mathbb{R}$$

satisfying $\phi'(x) = -x\,\phi(x)$

Let $f(x) = f(0) + \int_0^x f'(x)\,dx$

be an absolutely continuous function.

If $E_P(f'(G)^2) < \infty$

then also $E_P(f(G)^2) < \infty$

$$\int_{\mathbb{R}} f(x)^2 \phi(x)\,dx = E(f(G)^2) =$$

$$= \int_{\mathbb{R}} \left(f(0) + \int_0^x f'(y)\,dy \right)^2 \phi(x)\,dx$$

by Cauchy Schwartz inequality it is enough to show that

2

$$\int_{\mathbb{R}} \left(\int_0^x f'(y)\,dy \right)^2 \phi(x)\,dx < \infty$$

)

$$\left(\text{since} \quad f(0)^2 \int_{\mathbb{R}} \phi(x)\,dx = f(0)^2 \right)$$

By Jensen inequality
applied to the uniform probability
measure on $[0, x]$

$$\int_{\mathbb{R}} \left(\int_0^x f'(y)\,dy \right)^2 \phi(x)\,dy$$

$$\leq \int_{\mathbb{R}} x \left(\int_0^x f'(y)^2\,dy \right) \phi(x)\,dx$$

by Gaussian integration
by parts

$$= -\int_0^\infty \left(\int_0^x f'(y)^2\,dy \right) \phi'(x)\,dx$$

$$+ \int_{-\infty}^0 \int_x^0 f'(y)^2 \phi'(x)\,dx$$

by Fubini

$$= \int_{0}^{+\infty} \left(\int_{+\infty}^{y} \phi'(x)\,dx \right) f'(y)^2\,dy$$

$$+ \int_{-\infty}^{0} \left(\int_{-\infty}^{y} \phi'(x)\,dx \right) f'(y)^2\,dy$$

$$= \int_{\mathbb{R}} \phi(y)\, f'(y)^2\,dy =$$

$$= E\left(f'(G)^2 \right) < \infty$$

by assumption

this shows

$$E\left(f(G)^2 \right) \le E\left(f'(G)^2 \right)$$