UH Probability Theory II, exercises 8, solutions (11.11.2015)

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

   Show that

   (a)

$$E_P(X) = \int_0^\infty P(X > t)dt = \int_0^\infty P(X \geq t)dt$$

   Hint: $t = \int\limits_0^\infty \mathbf{1}(s \leq t)ds$, use Fubini

   (b)

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

   Solution

$$E_P(X^n) = \int_0^\infty s^n P_X(ds) = \int_0^\infty \left(\int_0^s nt^{n-1}dt\right)P_X(ds)$$
$$= \int_0^\infty \left(\int_0^\infty \mathbf{1}(t \leq s)nt^{n-1}dt\right)P_X(ds)$$
$$\overset{(Fubini)}{=} n\int_0^\infty \left(\int_0^\infty \mathbf{1}(t \leq s)P_X(ds)\right)t^{n-1}dt = n\int_0^\infty \mathbb{P}(X \geq t)\Big)t^{n-1}dt =$$
$$n\int_0^\infty \mathbb{P}(X > t)\Big)t^{n-1}dt$$

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

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

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

$$|p(x)| \leq C_1 + C_2\exp(\lambda x^2/2)$$

   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) Use symmetry to show that $\forall n \in \mathbb{N}$ we have $E_P(G^{2n+1}) = 0$ for all the odd moments.

Solution

$$\int_R x^{2n+1}\phi(x)dx = \int_{-\infty}^{0} x^{2n+1}\phi(x)dx + \int_0^{\infty} x^{2n+1}\phi(x)dx$$
$$= \int_{-\infty}^{0} x^{2n+1}\phi(-x)dx + \int_0^{\infty} x^{2n+1}\phi(x)dx$$
$$= -\int_0^{\infty} x^{2n+1}\phi(x)dx + \int_0^{\infty} x^{2n+1}\phi(x)dx = 0$$

where $\int_0^\infty x^{2n+1}dx < \infty$ since $E_P(|G|^{2n+1}) < \infty$.

(b) Compute $E_P(G^2)$.

Hint You can use the Gaussian integration by parts formula $E_P(f(G)G) = E_P(f'(G))$ after checking the integrability condition. Equivalently you can use the property of the standard Gaussian density $\partial_x\phi(x) = -\phi(x)x$ and use the usual integration by parts formula.

Solution For $f(x) = x$ with $f'(x) = 1$ we obtain $E_P(G^2) = E_P(GG) = E_P(f(G)G) = E_P(f\prime(G)) = E_P(1) = 1$.

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

Solution For $f(x) = x^{2n-1}$ with $f\prime(x) = (2n-1)x^{2n-2}$,

$E_P(G^{2n}) = E_P(G^{2n-1}G) = E_P(f(G)G) = E_P(f\prime(G)) = (2n-1)E_P(G^{2(n-1)}) = (2n-1)(2n-3)E_P(G^{2(n-2)}) = \dots$
$= (2n-1) \times (2n-3) \times (2n-5) \times \cdots \times 7 \times 5 \times 3 \times 1 := (2n-1)$

3. For $t \in \mathbb{R}$ compute the expectations:

(a) $E_P\big(G\mathbf{1}(G > t)\big)$
(b) $E_P\big(G\mathbf{1}(G \le t)\big)$
(c) $E_P\big(G^2\mathbf{1}(G > t)\big)$
(d) $E_P\big(G^2\mathbf{1}(G \le t)\big)$
(e) $E_P\big(G^3\mathbf{1}(G > t)\big)$
(f) $E_P\big(G^3\mathbf{1}(G \le t)\big)$

Hints: Show you can use the Gaussian integration by part formula $E_P(f(G)G) = E_P(f'(G))$ with $f(x) = \mathbf{1}(x > t)$. In this case $f'(x) = \delta_t(x) = \delta_0(x-t)$ is not a function but a generalized 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 $\{t\}$.

In order show that the integration by parts formula is correct also in this case, approximate the indicator $f(x) = \mathbf{1}(x > t)$ by the sequence $f_n(x) = \big((x-t)^+ n\big) \wedge 1$ which satisfies $0 \le f_n(x) \le f(x) \le 1\ \forall x$, and it is piecewise linear with derivative $f_n'(x) = n\mathbf{1}\big(t < x \le t + 1/n\big)$.

Apply the Gaussian integration by parts to $f_n(x)$ and use the dominated convergence Theorem to take limits.

Solution a) Formally for $f(x) = \mathbf{1}(x > t)$ with $f'(x) = \delta_t(x)$

$$E_P(G\mathbf{1}(G > t)) = E_P(Gf(G)) = E_P(f\prime(G)) = E_P(\delta_t(G)) = \int_{\mathbb{R}} \delta_t(x)\phi(x)dx$$

$$= \phi(t) = \frac{1}{\sqrt{2\pi}}\exp\big(-\frac{t^2}{2}\big)$$

However this is a bit suspicious, since $\delta_t(x)$ is not a function (it is a generalized function in the sense of distribution). We show that the formula is correct indeed. We just apply the Gaussian integration by parts to $f_n(x)$.

$$E_P(Gf_n(G)) = E_P(f_n'(G)) = nP\big(G \in (t, t+1/n]\big) = n\int_t^{t+1/n} \phi(x)dx$$

Note that $\forall\omega \in \Omega$, $G(\omega)f_n(G(\omega)) \to G(\omega)f(G(\omega))$ and $|G(\omega)f_n(G(\omega)) \le |G(\omega)| \in L^1(P)$ and by Lebesgue dominated convergence Theorem it follows that

$$E_P(Gf_n(G)) \to E_P(Gf(G))$$

on the other hands, since $x \mapsto \phi(x)$ is continuous,

$$\lim_{n\to\infty} E_P(f_n'(G)) = \lim_{n\to\infty} n\int_t^{t+1/n} \phi(x)dx = \phi(t) = E_P(\delta_t(G))$$

b) $E_P\big(G\mathbf{1}(G \le t)\big) = E_P(G) - E_P\big(G\mathbf{1}(G > t)\big) = 0 - \phi(t)$

c) $E_P\big(G^2\mathbf{1}(G > t)\big) = E_P\big(GG\mathbf{1}(G > t)\big) = E_P\big(Gf(G)\big) = E_P\big(f'(G)\big)$
where $f(x) = x\mathbf{1}(x > t)$ with derivative $\mathbf{1}(x > t) + x\delta_t(x)$. giving

$$E_P\big(G^2\mathbf{1}(G > t)\big) = E_P(\mathbf{1}(x > t)) + E_P\big(x\delta_t(x)\big) = 1 - \Phi(t) + t\phi(t) = \Phi(-t) + t\phi(t)$$

where $\Phi(t) = \int_{\infty}^{t} \phi(x)dx = P(G \le t)$ is the cumulative distribution function of the Gaussian distribution. By symmetry $P(G > t) = 1 - \Phi(t) = P(G \le -t) = \Phi(-t)$.

d)

$$E_P\big(G^2\mathbf{1}(G \le t)\big) = E_P(G^2) - E_P\big(G^2\mathbf{1}(G > t)\big) = \Phi(t) - t\phi(t)$$

e)

$$E_P\big(G^3\mathbf{1}(G > t)\big) = E_P\big(GG^2\mathbf{1}(G > t)\big) = E_P(f'(G))$$

where $f(x) = x^2\mathbf{1}(x > t)$ with derivative $2x\mathbf{1}(x > t) + x^2\delta_t(x)$

$$= 2E_P(G\mathbf{1}(G > t)) + E_P(G^2\delta_t(G)) = (2 + t^2)\phi(t)$$

f)

$$E_P\big(G^3\mathbf{1}(G \le t)\big) = E_P(G^3) - E_P\big(G^2\mathbf{1}(G > t)\big) = -(2 + t^2)\phi(t)$$

4. (χ^2_n 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

   $$X(\omega) = G_1(\omega)^2 + \cdots + G_n(\omega)^2, \qquad \text{and } \Pi_k(\omega) = \frac{G_k(\omega)^2}{X_n(\omega)}, \quad 1 \le k \le n$$

   Note that $\Pi_k(\omega) \in [0, 1]$ and $\sum_{k=1}^{n} \Pi_k(\omega) = 1$, so that
   This means that for each ω, the random vector $\Pi(\omega) = (\Pi_1(\omega), \dots, \Pi_n(\omega))$ belongs to the simplex

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

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

Use the change of variable formula for the bijection $f : \mathbb{R}_+^n \to \mathbb{R}_+^n$ defined as

$$g = \big(g_1, \dots, g_n\big) \mapsto f(g) = \left(\sum_{k=1}^n g_k^2, \frac{g_2^2}{\sum_{k=1}^n g_k^2}, \dots, \frac{g_n^2}{\sum_{k=1}^n g_k^2} \right)$$

to show that X and Π are independent, computing also the probability density of X on $\mathbb{R}^+$ and probability the density of $(\Pi_2, \dots, \Pi_n)$ on $[0,1]^{n-1}$.

By the way, the distribution of X is called chi-square 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.

Solution We do this in two stages.

Let $Y_k(\omega) = G_k(\omega)^2$. For $t \ge 0$

$$P(Y_k \le t) = P(G_k^2 \le t) = P(|G_k| \le \sqrt{t}) = P(-\sqrt{t} \le G_k \le \sqrt{t}) = 2\Phi(\sqrt{t}) - 1,$$

where $\Phi(t) = P(G \le t)$ is the cumulative distribution function of G, and the probability density function of $Y_k(\omega)$ is given by

$$\frac{\partial}{\partial t} P(Y_k \le t) = 2\frac{\partial}{\partial t}\Phi(\sqrt{t}) = \phi(\sqrt{t})t^{-1/2} = \frac{1}{\sqrt{2\pi}} \exp(-t/2)t^{-1/2}$$

This is the probability density of the so called χ_1^2 chi-square distribution with 1 degree of freedom.

Since $G_1, \dots, G_n$ are P-independent implies that also $Y_1, \dots, Y_n$ are P-independent.

Now $X(\omega) = Y_1(\omega) + \dots + Y_n(\omega)$ and $\Pi_k(\omega) = Y_k(\omega)/X(\omega)$

The map

$$\Psi(y_1, \dots, y_n) = \left(x = y_1 + \dots + y_n, p_2 = \frac{y_2}{y_1 + \dots + y_n}, \dots, p_n = \frac{y_n}{y_1 + \dots + y_n} \right)$$

is a diffeomorphism between $(0, \infty)^n$ and the open set

$$(0, +\infty) \times \big\{ (p_2, \dots, p_n) \in [0,1]^{n-1} : (p_2 + \dots + p_n) < 1 \big\}$$

.

$$Jf(g) = \left[\frac{\partial \psi_i}{\partial y_j} \right]_{ij} =$$

$$\frac{1}{(y_1 + \dots + y_n)} \begin{pmatrix} 0 & 0 \\ 0 & Id_{(n-1)\times(n-1)} \end{pmatrix} - \left(1, \frac{y_2}{(y_1 + \dots + y_n)}, \dots, \frac{y_n}{(y_1 + \dots + y_n)}\right)^\top (1, 1, \dots, 1)$$

by using the multilinearity and alternating properties of the determinant, it follows that

$$\det(Jf(g)) = (y_1 + \cdots + y_n)^{-(n-1)}$$

We obtain the joint density of $(X_1, \Pi_2, \ldots, \Pi_n)$

$$(2\pi)^{-n/2} \prod_{k=1}^{n} \exp\left(-\frac{y_k}{2}\right) y_1^{-1/2} y_2^{-1/2} \ldots y_n^{-1/2} \left| \det(J\psi(y)) \right|^{-1}$$

$$= (2\pi)^{-n/2} \exp\Bigl(-\frac{1}{2}\sum_{k=1}^{d} y_k\Bigr) \Bigl(\frac{y_1}{x}\Bigr)^{-1/2} \ldots \Bigl(\frac{y_n}{x}\Bigr)^{-1/2} x^{-n/2} \times x^{d-1} =$$

$$(2\pi)^{-n/2} \exp(-x/2) x^{n/2-1} \Bigl(1 - (p_2 + \cdots + p_n)\Bigr)^{-1/2} p_2^{1/2} \ldots p_n^{-1/2}$$

with $p_k > 0$ and $0 \le p_2 + \cdots + p_n \le 1$, and $p_1 = 1 - (p_2 + \cdots + p_n)$, and $(y_1, \ldots, y_n) = \Psi^{-1}(x, p_2, \ldots, p_n)$.

$$\int_0^\infty \exp(-x/2) x^{n/2-1} dx = 2^{n/2} \int_0^\infty \exp(-u) u^{n/2-1} du = 2^{n/2} \Gamma(n/2)$$

where the $\Gamma(z)$ is the Gamma function defined by the integral. Therefore we have the factorization

$$\frac{1}{2^{n/2}\Gamma(n/2)} \exp(-x/2) x^{n/2-1} \times \frac{\Gamma(n/2)}{\Gamma(1/2)^n} \Bigl(1 - (p_2 + \cdots + p_n)\Bigr)^{-1/2} p_2^{-1/2} \ldots p_n^{-1/2}$$

where $\Gamma(1/2) = \sqrt{\pi}$. This means that X is χ_n^2 (chi-square distributed with n-degrees of freedom) while $(\Pi_1, \ldots, \Pi_n)$ is the Dirichlet distributed with parameters $(1/2, \ldots, 1/2)$ taking values in the simplex Δ_n. Note that the distribution of $(\Pi_1, \Pi_2, \ldots, \Pi_n)$ is degenerate, since $\Pi_1 = 1 - (\Pi_2 + \cdots + \Pi_n)$, it does not have a density with respect to the n-dimensional Lebesgue measure. However $(\Pi_2, \ldots, \Pi_n)$ has density with respect to the $(n-1)$-dimensional Lebesgue measure.