UH Malliavin Calculus, Fall 2016, Exercises 1-2 (14 and 21.9 2016)

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

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

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

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

which implies that $E_P(|G|^p) < \infty$ and $G \in L^p(P)$ for all exponents $p > 0$. Similarly all exponential moments $E_P\big(\exp(tG)\big) = \exp(t^2/2)$ are finite $\forall t \in \mathbb{R}$. 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.

(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.

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

2. 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) $E_P\big(G^3\mathbf{1}(G \le t)\big)$
(g) $E_P\big(G^4\mathbf{1}(G > t)\big)$
(h) $E_P\big(G^4\mathbf{1}(G \le t)\big)$

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.

3. Let $H \subset L^2(\Omega, \mathcal{F}, P)$ be a closed (with respect to the L^2 norm) subspace of random variables, and $X \in L^2(P)$. There exists an element $Y \in H$ which minimizing the L^2 distance from X among all H elements.

$$E_P\big((Y-X)^2\big) \le E_P\big((V-X)^2\big) \quad \forall V \in H$$

It is also true that $(Y - X) \perp H$, ($(X - Y)$ is orthogonal to H), equivalently

$$E_P(YV) = E_P(XV) \quad \forall V \in H$$

We denote $\Pi_H X = Y$ as the orthogonal projection of X into the subspace H.

For example if $\mathcal{G} \subseteq \mathcal{F}$ is a sub σ-algebra, $H = L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of $L^2(\Omega, \mathcal{G}, P)$ and in this case $\Pi_H X$ coincides with the conditional expectation $E_P(X|\mathcal{G})$.

(a) Show that the L^2-projection Π_H is a linear operator: when $X, Z \in L^2(P),\ a, b \in \mathbb{R}$,

$$\Pi_H(aX + bZ) = a\Pi_H X + b\Pi_H Z,$$

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

(c) Show that the projection does not increase the L^2 norm:

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

These properties characterize projection operators.

The Next exercises are about combining the idea of taking $L^2(P)$-projections to the linear span subspace of $L^2(P)$ random variable, together with the integration by parts formula for Gaussian, Poisson, and Bernoulli variables.

4. Let $G(\omega) \sim N(0,1)$ be a standard Gaussian variable with probability density $\phi(y) = (2\pi)^{-1/2} \exp\big(-y^2/2\big)$, and let $f(x)$ be a differentiable function with $E_P(f(G)^2) < \infty$ and $E_P(|f'(G)|) < \infty$.

   (a) Use the Gaussian integration by parts formula to show that

   $$\widehat{f}(G) = E_P(f(G)) + E_P(f'(G))G(\omega)$$

   is the best linear approximation of $f(G)$ based in the closed linear span of $\{\ 1, G(\omega)\}$ in least square sense, meaning that $\widehat{a} = E_P(f(G))$ and $\widehat{b} = E_P(f'(G))$ are minimizing the mean square error

   $$E_P\big(\big\{f(G) - (a + bG)\big\}^2\big), \quad a, b \in \mathbb{R}$$

   (b) Now we consider the same linear approximation in the multivariate case. where we use the following extension of the linear predictor formula from Example 9.1.1. in the lecture notes:
   When $X(\omega) = (X_1(\omega), \ldots, X_T(\omega)) \in L^2(\Omega, \mathcal{F}, P)$, then the following multivariate formula holds: for $Y = (Y_1, \ldots, Y_d)$ is another random variable in $L^2(P)$,

   $$\widehat{Y} = E_P(Y) + (X - E_P(X))\mathrm{Cov}(Y, Y)^{-1}\mathrm{Cov}(X, Y)$$

   where M^{-1} denoted the inverse of a matrix M and $\mathrm{Cov}(X, Y)_{ij} = E(X_i Y_j) - E(X_i)E(X_j)$ is the covariance between X_i and Y_j,
   and $\widehat{Y_i}$ is the $L^2(P)$-projection of Y_i to the linear span of $\{1, X_1, \ldots, X_d\}$.
   Let $G(\omega) = (G_1(\omega), \ldots, G_T(\omega)) \in \mathbb{R}^T$ 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 $E_P\big(f(G_1, \ldots G_n)^2\big) < \infty$ and

   $$E_P\left(\left|\frac{\partial}{\partial x_t} f(G_1, \ldots, G_T)\right|\right) < \infty$$

Show that

$$\widehat{f}(G_1,\dots,G_T) = E_P(f(G_1,\dots,G_T)) + \sum_{t=1}^{T} E_P\left(\frac{\partial}{\partial x_t} f(G_1,\dots,G_T)\right) G_t$$

is the best linear 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

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

(c) Next we consider the correlated case: let $A = (A_{st})$ be a non-singular $T \times T$ matrix, $G = (G_1,\dots,G_T)$ with i.i.d. standard Gaussian coordinates as before and let $X = (X_1,\dots,X_T) = AG^\top$ with coordinates

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

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

$$E_P(f(X_1,\dots,X_T)^2) < \infty$$

and

$$E_P\left(\left|\frac{\partial}{\partial x_t} f(X_1,\dots,X_T)\right|\right) < \infty$$

Compute the coefficients of the best linear approximation $\widehat{f}(X_1,\dots,X_T)$ of $f(X_1,\dots,X_T)$ in the linear span of $\{1, X_1,\dots,X_T\}$ minimizing the mean square error

$$E_P\left(\left\{ f(X_1,\dots,X_T) - \left(c_0 + \sum_{t=1}^{T} c_t X_t\right)\right\}^2\right)$$

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

$$P_\lambda\big(N = k\big) = \exp(-\lambda)\frac{\lambda^k}{k!} \quad \text{for } k \in \mathbb{N}\,.$$

and $(f(k) : k \in \mathbb{N})$ a sequence with $E(f(N)^2) < \infty$.

(a) Prove the Stein equation or integration by parts formula for Poisson-λ random variables:

$$E_\lambda\big(f(N)N\big) = \lambda E_\lambda\big(f(N+1)\big)$$

(b) Use the integration by parts formula for Poisson random variable to compute the first Poisson moments. $E_\lambda(N) = \lambda$, $E_\lambda(N^2) = \lambda^2 + \lambda$.

(c) Show that

$$\widehat{f}(N) = E_\lambda(f(N)) + E_\lambda(f(N+1) - f(N))(N - \lambda)$$

is the best linear estimator of $f(N)$ depending on N on the closed linear span of $\{1, N(\omega)\}$, with coefficients minimizing the mean square error

$$E_P(\{f(N) - (a + bN)\}^2)$$

(d) Let now Let $N(\omega) = (N_1(\omega), \dots, N_T(\omega)) \in \mathbb{N}^T$ where the coordinates are $N_t(\omega)$ are independent and Poisson(λ_t) distributed for $t = 1, \dots, T$, respectively, with $\lambda_t > 0$ (possibly different).
Let $f : \mathbb{N}^T \to [0, +\infty)$ be a function with $E_P(f(N_1, \dots, N_T)^2) < \infty$.
Show that

$$\widehat{f}(N_1, \dots, N_T) = E_P(f(N_1, \dots, N_T)) +$$
$$\sum_{t=1}^{T} E_P\Big(f(N_1, \dots, N_{t-1}, 1 + N_t, N_{t+1} \dots, N_T) - f(N_1, \dots, N_{t-1}, N_t, N_{t+1} \dots,, N_T)\Big)(N_t - \lambda_t)$$

is the best linear approximation of $f(N_1, \dots, N_T)$ in the linear span of $\{1, N_1, \dots, N_T\}$ with coefficients $c_t \in \mathbb{R}$ minimizing the mean square error

$$E_P\left(\left\{f(N_1, \dots, N_T) - \left(c_0 + \sum_{t=1}^{T} c_t N_t\right)\right\}^2\right)$$

6. Let $G(\omega)$ be a standard Gaussian random variables.

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

(a) Use the Gaussian integration by parts formula 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: when $E_P(f(G)^2) < \infty$ and $E_P(\partial f(G)^2) < \infty$, and for another differantiable h with $E_P(\partial h(G)^2) < \infty$,

$$E_P(h(G)\partial f(G)) = E_P(f(G)\partial^* h(G))$$

$$\partial^* f(x) := xf(x) - \partial f(x)$$

is the adjoint of the derivative operator ∂ in the space $L^2(\mathbb{R}, \mathcal{F}, \phi(x)dx)$, 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 $h_n(x) = (\partial^{*n}1)(x) = \partial^{*n} h_{n-1}(x)$.
Compute the first five $h_n(x)$ Hermite polynomials for $n = 1, 2, 3, 4, 5$.

(c) Show that $E(h_n(G)) = 0$

(d) Show that $E(h_n(G), h_m(G)) = n!\delta_{nm}$ and in particular the random variables $h_n(G)$ and $h_m(G)$ are orthogonal in $L^2(\Omega, \mathcal{F}, P)$ Hint: use extended Gaussian integration by parts, and that ∂^* is the adjoint of the derivative in $L^2(\mathbb{R}, \mathcal{F}, \phi(x)dx)$.

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

$$\widehat{f}(G) = E_P(G) + \sum_{k=1}^{n} \frac{E_P(\partial^k f(G))}{k!} h_k(G) = E_P(G) + \sum_{k=1}^{n} \frac{E_P(f(G)h_k(G))}{k!} h_k(G)$$

is the best polynomial approximation of $f(G)$ in the linear span of $\{h_0(G) = 1, h_1(G) = G, \dots, h_n(G)\}$ with coefficients minimizing the least square error

$$E_P\left(\left\{f(G) - \left(\sum_{k=0}^{n} c_k h_n(G)\right)\right\}^2\right)$$

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

8. We compute linear projections with Bernoulli random variables. Let $X(\omega)$ be a binary random variable with

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

and p in $[0, 1]$.

(a) Show that then best linear approximation of $f(X)$ for $f : \{0, 1\} \to \mathbb{R}$ in the linear span of $\{1, X_1(\omega)\}$ in mean square sense is given by

$$\widehat{f}(X) = E_p(f(X)) + (f(1) - f(0))(X - p)$$

where $E_P(X) = E_P(X^2) = p$.

(b) Actually in this case the approximation is exact: check that $\widehat{f}(X) = f(X)$!

(c) For $X_1(\omega), \ldots, X_T(\omega)$ independent random variables with

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

and p_t in $[0, 1]$, and $f : \{0, 1\}^T \to \mathbb{R}$, show that best linear approximation of $f(X)$ in the linear span of $\{1, X_1(\omega), \ldots, X_T(\omega)\}$ in mean square sense is given

$$\widehat{f}(X_1, \ldots, X_T) = E_P(f(X)) +$$
$$\sum_{t=1}^{T} E_P\big(f(X_1, \ldots, X_{t-1}, 1, X_{t+1}, \ldots X_{t_n}) - f(X_1, \ldots, X_{t-1}, 0, X_{t+1}, \ldots X_{t_n})\big)(X_t(\omega) - p_t)$$