Name: Tuan Pham
ID: 4652218

Math 8652: Theory of Probability

Homework #4

① Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_N$ be σ-subfields of $\mathcal{F}$, and $\xi_1, \xi_2, \dots, \xi_N$ be $[0,\infty)$-valued random variables satisfying

- ξ_n is $\mathcal{F}_n$-measurable $\quad \forall 1 \leq n \leq N$,
- $E|\xi_n| < \infty \quad \forall 1 \leq n \leq N$.

First, suppose $(\xi_1, \dots, \xi_N)$ is a submartingale with respect to $(\mathcal{F}_1, \dots, \mathcal{F}_n)$. We show that there exist random variables $\eta_1, \eta_2, \dots, \eta_N$ such that

- $0 \leq \eta_1 \leq \eta_2 \leq \dots \leq \eta_N = \xi_N$,
- $\xi_n = E(\eta_n | \mathcal{F}_n) \quad \forall 1 \leq n \leq N$.

Define $\zeta_n = \dfrac{\xi_n}{E(\xi_{n+1}|\mathcal{F}_n)}$ with the convention $\frac{0}{0} = 0$. Because $0 \leq \xi_n \leq E(\xi_{n+1}|\mathcal{F}_n)$, $0 \leq \zeta_n \leq 1$. Because ξ_n and $E(\xi_{n+1}|\mathcal{F}_n)$ are $\mathcal{F}_n$-measurable, ζ_n is also $\mathcal{F}$-measurable. Then

$$\xi_n = \frac{\xi_n}{E(\xi_{n+1}|\mathcal{F}_n)} E(\xi_{n+1}|\mathcal{F}_n) = \zeta_n E(\xi_{n+1}|\mathcal{F}_n) = E(\zeta_n \xi_{n+1} | \mathcal{F}_n). \qquad (1)$$

Take $m \in \{1, 2, \dots, N-1\}$. We now show ~~that~~ by induction in $k \in \{1, 2, \dots, N-m\}$ that

$$\xi_m = E(\zeta_m \zeta_{m+1} \cdots \zeta_{m+k-1} \xi_{m+k} | \mathcal{F}_m). \qquad (2)$$

For $k=1$, (2) is true because of (1). Suppose (2) is true for some $k \in \{1, 2, \dots, N-m-1\}$.

Applying (1) for $n = m+k$, we get $\xi_{m+k} = E(\zeta_{m+k} \xi_{m+k+1} | \mathcal{F}_{m+k})$. Because $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n$, the random variables $\zeta_m, \zeta_{m+1}, \dots, \zeta_{m+k-1}$ are $\mathcal{F}_{m+k}$-measurable.

Substituting ξ_{m+k} above into (2), we get

$$\begin{aligned}\xi_m &= E(\zeta_m \zeta_{m+1} \cdots \zeta_{m+k-1} E(\zeta_{m+k} \xi_{m+k+1} | \mathcal{F}_{m+k}) | \mathcal{F}_m) \\ &= E(E(\zeta_m \zeta_{m+1} \cdots \zeta_{m+k-1} \zeta_{m+k} \xi_{m+k+1} | \mathcal{F}_{m+k}) | \mathcal{F}_m) \\ &= E(\zeta_m \zeta_{m+1} \cdots \zeta_{m+k} \xi_{m+k+1} | \mathcal{F}_m).\end{aligned}$$

Thus, (2) is also true for $k+1$.

In (2), we take $k = N-m$. Then $\xi_m = E(\zeta_m \zeta_{m+1} \cdots \zeta_{N-1} \xi_N | \mathcal{F}_m)$. Define

$$\eta_N = \xi_N,$$

$$\eta_m = \zeta_m \zeta_{m+1} \cdots \zeta_{N-1} \xi_N \quad \forall 1 \leq m \leq N-1.$$

Then $\eta_1, \eta_2, \dots, \eta_N$ are $[0,\infty)$-valued random variables and $\xi_m = E(\eta_m | \mathcal{F}_m)$. Also,

$$\eta_m = \zeta_m \eta_{m+1} \leq \eta_{m+1} \quad \forall 1 \leq m \leq N-1.$$

Conversely, suppose there exist random variables $\eta_1, \eta_2, \dots, \eta_N$ such that

- $0 \leq \eta_1 \leq \eta_2 \leq \dots \leq \eta_N = \xi_N$,
- $\xi_n = E(\eta_n | \mathcal{F}_n) \quad \forall 1 \leq n \leq N$.

We show that $(\xi_1, \dots, \xi_N)$ is a submartingale with respect to $(\mathcal{F}_1, \dots, \mathcal{F}_n)$. For $1 \leq n < m \leq N$,

$$E(\xi_m | \mathcal{F}_n) = E(E(\eta_m | \mathcal{F}_m) | \mathcal{F}_n) = E(\eta_m | \mathcal{F}_n) \geq E(\eta_n | \mathcal{F}_n) = \xi_n.$$

(2) Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(X_1, X_2, \dots, X_N)$ be an independent vector of random variables, each having standard normal distribution. Put $w_N = X_1 + X_2 + \dots + X_N$. For each $A \in \mathcal{F}$, define

$$P'(A) = E I_A \exp\left(w_N - \frac{N}{2}\right).$$

First, we show that P' is a probability measure on $\mathcal{F}$. Let $A_1, A_2, A_3, \dots$ be a sequence of pairwise disjoint events. Then

$$P'\left(\bigcup_{n=1}^{\infty} A_n\right) = E\, I_{\cup A_n} \exp\left(w_N - \frac{N}{2}\right) = E\left(\sum_{n=1}^{\infty} I_{A_n} \exp\left(w_N - \frac{N}{2}\right)\right)$$

$$= \sum_{n=1}^{\infty} E I_{A_n} \exp\left(w_N - \frac{N}{2}\right) \quad \text{(Monotone Convergence theorem)}$$

$$= \sum_{n=1}^{\infty} P'(A_n).$$

Now we compute $P'(\Omega)$.

$$P'(\Omega) = E \exp\left(w_N - \frac{N}{2}\right) = E \prod_{k=1}^{N} \exp\left(X_k - \frac{1}{2}\right)$$

$$= \prod_{k=1}^{N} E \exp\left(X_k - \frac{1}{2}\right) \quad \text{(because } X_1, \dots, X_N \text{ are independent)}$$

$$= \left\{ E \exp\left(X_1 - \frac{1}{2}\right)\right\}^N \quad \text{(because } X_1, \dots, X_N \text{ have the same distribution)}$$

$$= \left\{ \exp\left(-\frac{1}{2}\right) E \exp(X_1) \right\}^N. \qquad (1)$$

Because X_1 has standard normal distribution, its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Then
$$E \exp(X_1) = \int_{\mathbb{R}} \exp(x) f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(x - \frac{x^2}{2}\right) dx$$

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

$$\overset{t=x-1}{=\!=} \exp\left(\frac{1}{2}\right) \underbrace{\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{2}\right) dt}_{=1}$$

$$= \exp\left(\frac{1}{2}\right). \qquad (2)$$

$\boxed{4}$

Then $\exp(-\frac{1}{2})\, E \exp(X_1) = 1$. Then (1) gives us $P'(\Omega) = 1$.

Next, put $Y_n = X_n - 1$ for all $1 \le n \le N$. Because $(X_1, \dots, X_N)$ is an independent vector, so is $(Y_1, \dots, Y_N)$. We now show that each Y_k, when viewed as a random variable on $(\Omega, \mathcal{F}, P')$, has standard normal distribution. That is to show

$$P'(Y_k \le x) = \int_{-\infty}^{x} f(t)\,dt \qquad \forall x \in \mathbb{R}.$$

We will only consider $k=1$. Other values of k are treated similarly (but the notations are not as convenient).

$$P'(Y_1 \le x) = P'(X_1 \le x+1) = E\, I_{X_1 \le x+1} \exp\left(w_N - \frac{N}{2}\right)$$

$$= E\left[I_{X_1 \le x+1} \exp(X_1 - \tfrac{1}{2}) \exp(X_2 - \tfrac{1}{2}) \cdots \exp(X_N - \tfrac{1}{2})\right]$$

$$= E\left[I_{X_1 \le x+1} \exp(X_1 - \tfrac{1}{2})\right] E\left[\exp(X_2 - \tfrac{1}{2})\right] \cdots E\left[\exp(X_N - \tfrac{1}{2})\right]. \qquad (3)$$

(because $X_1, X_2, \dots, X_N$ are independent)

Since $X_1, X_2, \dots, X_N$ have the same distribution,

$$E\left[\exp(X_2 - \tfrac{1}{2})\right] = \cdots = E\left[\exp(X_N - \tfrac{1}{2})\right] = E\left[\exp(X_1 - \tfrac{1}{2})\right] \overset{(2)}{=} 1.$$

Then (3) reduces to

$$P'(Y_1 \le x) = E\left[I_{X_1 \le x+1} \exp(X_1 - \tfrac{1}{2})\right]$$

$$= \int_{\mathbb{R}} I_{t \le x+1} \exp(t - \tfrac{1}{2}) f(t)\,dt$$

$$= \int_{-\infty}^{x+1} \exp(t - \tfrac{1}{2}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x+1} \exp\left(-\frac{(t-1)^2}{2}\right) dt$$

$$\overset{s=t-1}{=\!=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{s^2}{2}\right) ds$$

$$= \int_{-\infty}^{x} f(s)\,ds.$$

(3) Let $(\xi, \xi_1, \dots, \xi_n)$ be a Gaussian vector in $\mathbb{R}^{n+1}$. Let $f: \mathbb{R} \to \mathbb{R}$ be a Borel function such that $E|f(\xi)| < \infty$. Put $m = E(\xi | \xi_1, \dots, \xi_n)$ and $\eta = \xi - m$.

We show that

- η has normal distribution $N(0, \sigma^2)$,
- $E(f(\xi) | \xi_1, \dots, \xi_n) = Ef(x+\eta)\big|_{x=m}$ a.s.

Define a map $F: \mathbb{R}^{n+1} \to \mathbb{R}$, $F(a, b_1, \dots, b_n) = E[\xi - (a + b_1\xi_1 + \dots + b_n\xi_n)]^2$.

Because F is a (multivariate) quadratic polynomial, it attains minimum at some point $(a^0, b_1^0, \dots, b_n^0) \in \mathbb{R}^{n+1}$. Then

$$0 = \frac{\partial F}{\partial a}(a^0, b_1^0, \dots, b_n^0) = -2E[\xi - (a^0 + b_1^0\xi_1 + \dots + b_n^0\xi_n)],$$

$$0 = \frac{\partial F}{\partial b_k}(a^0, b_1^0, \dots, b_n^0) = -2E\xi_k[\xi - (a^0 + b_1^0\xi_1 + \dots + b_n^0\xi_n)] \quad \forall 1 \le k \le n.$$

Put $\tilde{m} = a^0 + b_1^0\xi_1 + \dots + b_n^0\xi_n$ and $\tilde{\eta} = \xi - \tilde{m}$. Then

$$\begin{cases} E\tilde{\eta} = 0, \\ E\tilde{\eta}\xi_k = 0 \quad \forall 1 \le k \le n. \end{cases}$$

Because $(\xi, \xi_1, \dots, \xi_n)$ is a normally distributed vector, $\xi - b_1^0\xi_1 - \dots - b_n^0\xi$ is a normally distributed random variable. Then $\tilde{\eta}$ is also normally distributed.

$$E[(\tilde\eta - E\tilde\eta)(\xi_k - E\xi_k)] = E\tilde\eta(\xi_k - E\xi_k) = E\tilde\eta\,\xi_k - E\tilde\eta\, E\xi_k = E\tilde\eta\,\xi_k = 0.$$

Thus, $\tilde\eta$ and ξ_k are uncorrelated. Put $\zeta = (\xi_1, \dots, \xi_n)$. We show that $(\tilde\eta, \zeta)$ is an independent pair. Because $(\xi, \xi_1, \dots, \xi_n)$ is a normally distributed vector, so is $(\tilde\eta, \xi_1, \dots, \xi_n)$. Theorem 20, Fristedt-Gray page 237, gives us the formula of the characteristic function of $(\tilde\eta, \xi_1, \dots, \xi_n)$. For $t_0, t_1, \dots, t_n \in \mathbb{R}$,

$$\beta_{(\tilde\eta,\zeta)}(t_0, t_1, \dots, t_n) = \exp\left(it_0\tilde\eta + i\sum_{k=1}^{n} \xi_k t_k - \frac{1}{2}(t_0, t_1, \dots, t_n) S (t_0, t_1, \dots, t_n)^T\right), \quad (1)$$

where S is the covariance matrix of $(\tilde\eta, \xi_1, \dots, \xi_n)$. Because $\mathrm{Cov}(\tilde\eta, \xi_k) = 0$, S is of block form

$$S = \begin{pmatrix} \sigma^2 & \\ & \boxed{\mathrm{Cov}(\zeta)} \end{pmatrix}$$

where $\sigma^2 = \mathrm{Var}\,\tilde\eta$. Then (1) becomes

$$\beta_{(\tilde\eta,\zeta)}(t_0, t_1, \dots, t_n) = \exp\left(it_0\tilde\eta + i\sum_{k=1}^{n} \xi_k t_k - \frac{\sigma^2}{2}t_0^2 - \frac{1}{2}(t_1, \dots, t_n)\mathrm{Cov}(\zeta)(t_1, \dots, t_n)^T\right)$$

$$= \exp\left(it_0\tilde\eta - \frac{\sigma^2}{2}t_0^2\right) \exp\left(i\sum_{k=1}^{n} \xi_k t_k - \frac{1}{2}(t_1, \dots, t_n)\mathrm{Cov}(\zeta)(t_1, \dots, t_n)^T\right)$$

$$= \beta_{\tilde\eta}(t_0)\,\beta_{\zeta}(t_1, \dots, t_n).$$

Let (η', ζ') be an independent pair such that η' has the same distribution as η, and ζ' has the same distribution as ζ. Then

$$\beta_{(\eta',\zeta')}(t_0, t_1, \dots, t_n) = \beta_{\eta'}(t_0)\,\beta_{\zeta'}(t_1, \dots, t_n) = \beta_{\tilde\eta}(t_0)\,\beta_{\zeta}(t_1, \dots, t_n) = \beta_{(\tilde\eta,\zeta)}(t_0, t_1, \dots, t_n).$$

Thus, (η', ζ') and $(\tilde{\eta}, \zeta)$ have the same distribution. Let $A \in \mathcal{B}(\mathbb{R})$ and $B \in \mathcal{B}(\mathbb{R}^n)$. Then

$$P((\tilde{\eta}, \zeta) \in A \times B) = P((\eta', \zeta') \in A \times B) = P(\eta' \in A) P(\zeta' \in B)$$
$$= P(\tilde{\eta} \in A)\, P(\zeta \in B).$$

This implies $\tilde{\eta}$ and ζ are independent.

Put $\mathcal{G} = \sigma(\zeta) = \sigma(\xi_1, \ldots, \xi_n)$. We have $E(\tilde{\eta} | \mathcal{G}) = E(\tilde{\eta}) = 0$; $E(\tilde{m} | \mathcal{G}) = \tilde{m}$ because $\tilde{m}$ is $\mathcal{G}$-measurable.

$$E(\xi | \mathcal{G}) = E(\tilde{\eta} + \tilde{m} | \mathcal{G}) = E(\tilde{\eta} | \mathcal{G}) + E(\tilde{m} | \mathcal{G}) = \tilde{m}.$$

Hence, $\tilde{m} = m$. Then $\eta = \xi - m = \xi - \tilde{m} = \tilde{\eta}$. Then η has distribution $N(0, \sigma^2)$.

Next, we show that $E(f(\xi) | \mathcal{G}) = Ef(x + \eta)|_{x=m}$. Define a function $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $h(x_1, x_2) = f(x_1 + x_2)$. This is a Borel function on $\mathbb{R}^2$. Then $f(\xi) = f(m + \eta) = h(m, \eta)$ and $E|h(m, \eta)| = E|f(\xi)| < \infty$. Write $h = h^+ - h^-$. By a theorem in lecture 02/18/2015,

$$E(h^+(m, \eta) | \mathcal{G}) = E h^+(x, \eta)|_{x=m},$$

$$E(h^-(m, \eta) | \mathcal{G}) = E h^-(x, \eta)|_{x=m}.$$

Subtracting the latter identity from the former, we get

$$E(h(m, \eta) | \mathcal{G}) = E h(x, \eta)|_{x=m}.$$

In other words, $E(f(\xi) | \mathcal{G}) = Ef(x + \eta)|_{x=m}$.

Consider the case $\sigma^2 > 0$. That is the case $E\eta^2 > 0$, i.e. ξ is not $\mathcal{G}$-measurable. Since η has distribution $N(0, \sigma^2)$,

$$E(f(\xi)|\mathcal{G}) = Ef(x+\eta)|_{x=m} = \int_{\mathbb{R}} f(x+t) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2}\right) dt \Big|_{x=m}$$

$$\overset{s=x+t}{=\!=\!=} \int_{\mathbb{R}} f(s) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(s-x)^2}{2}\right) ds \Big|_{x=m}$$

$$= \int_{\mathbb{R}} f(s) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(s-m)^2}{2}\right) ds.$$

④ Let $(\Omega, \mathcal{F}, P)$ be a probability space, X be a real-valued random variable with $EX^2 < \infty$, and $\mathcal{G}$ be a σ-subfield of $\mathcal{F}$. We show that

$$EX^2 = E[E(X|\mathcal{G})]^2 + E[X - E(X|\mathcal{G})]^2.$$

We know that $L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$. Put $Y = E(X|\mathcal{G}) \in L^2(\Omega, \mathcal{G}, P)$. We show that Y is the orthogonal projection of X on $L^2(\Omega, \mathcal{G}, P)$. For each $Z \in L^2(\Omega, \mathcal{G}, P)$,

$$(Y, Z) = E(YZ) = E(Z\, E(X|\mathcal{G})) = E(E(ZX|\mathcal{G})) = E(ZX) = (X, Z).$$

Thus, $$(X - Y, Z) = 0 \quad \forall Z \in L^2(\Omega, \mathcal{G}, P).$$

Therefore, Y is the orthogonal projection of X on $L^2(\Omega, \mathcal{G}, P)$. Then

$$(X, X) = (X-Y, X-Y) + 2(X, Y) - (Y, Y) = (X-Y, X-Y) + 2\underbrace{(X-Y, Y)}_{=0} + (Y, Y).$$

Thus, $$EX^2 = E(X-Y)^2 + EY^2$$

$$= E[X - E(X|\mathcal{G})]^2 + E[E(X|\mathcal{G})]^2.$$