2. Pogojne pričakovane vrednosti

2.1. Obstoj in lastnosti

Če je X slučajna spremenljivka z vrednostmi v $\mathbb{R}$, je družina $\{X^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}$ σ-algebra, ki je vsebovana v $\mathcal{F}$ (X je slučajna spremenljivka na $(\Omega, \mathcal{F}, P)$). Očitno je ta σ-algebra najmanjša, glede na katero je X merljiva. Označimo jo z $\sigma(X)$. Isti razmislek velja za več slučajnih spremenljivk. Pripadajočo σ-algebro označimo z $\sigma(X_1, X_2, \ldots, X_n)$.

Pri dodiplomski Verjetnosti smo za diskretni slučajni spremenljivki X in Y definirali

$$E(Y \mid X = x) = \sum_y y\, P(Y = y \mid X = x).$$

Pri Statistiki smo rekli

$E(Y \mid X = x) = \psi(x)$ in definirali

$E(Y \mid X) = \psi(X)$. Slučajna spremenljivka $E(Y \mid X)$ je na množicah $\{X = x\}$ konstantna in enaka

$$\frac{E[Y \cdot 1(X = x)]}{P(X = x)}.$$

Če je X' druga slučajna spremenljivka, ki generira enako particijo $\{X = x\}$ kot X, je očitno $E(Y \mid X) = E(Y \mid X')$.

Pogojna pričakovana vrednost je odvisna samo od particije, ki jo generira X. Kaj pa to particijo povzema? $\sigma(X)$! Ideja je, da definiramo pogojne pričakovane vrednosti ne glede na X, ampak glede na poljubno σ-algebro $\mathcal{G} \subseteq \mathcal{F}$. V diskretnem primeru za $G \in \sigma(X)$ velja

$$E[\psi(X) \cdot 1_G]$$

$$= E[E(Y|X) \cdot 1_G]$$

$$= E[E(Y|X) \sum_{k \in K} 1_{G_k}]$$

$$= \sum_{k \in K} \frac{E[Y \cdot 1_{G_k}]}{P(G_k)} \cdot P(G_k)$$

$$= E(Y \cdot 1_G).$$

To je motivacija za splošno definicijo

Definicija: Naj bo X slučajna spremenljivka z vrednostmi v $\mathbb{R}$ na $(\Omega, \mathcal{F}, P)$. Naj bo $E|X| < \infty$ in $\mathcal{G} \subseteq \mathcal{F}$ σ-algebra. Pogojna pričakovana vrednost $E(X \mid \mathcal{G})$ je slučajna spremenljivka z lastnostmi:

(i) $E(X \mid \mathcal{G})$ je $\mathcal{G}$-merljiva.

(ii) Za vsak $G \in \mathcal{G}$ velja

$$E[E(X \mid \mathcal{G}) \cdot 1_G]$$

$$= E[X \cdot 1_G].$$

Kako je z obstojem in enoličnostjo?

Enoličnost je preprosta. Če sta $\tilde{X}$ in $\tilde{X}'$ $\mathcal{G}$-merljivi in ustrezata definiciji pogojne pričakovane vrednosti, je

$$E[\tilde{X} \cdot 1_{\mathcal{G}}] = E[\tilde{X}' 1_{\mathcal{G}}].$$

Vzemimo $G = \{\tilde{X} > \tilde{X}'\}$. Velja

$$E[(\tilde{X} - \tilde{X}') 1_G] = 0,$$

ampak razlika $\tilde{X} - \tilde{X}'$ je na G strogo pozitivna, zato je $P(G) = 0$. Vlogi lahko obrnemo. Sledi, da je $P(\tilde{X} \neq \tilde{X}') = 0$.

Obstoj je bolj komplicirano vprašanje.

Iz teorije mere bomo potegnili še en izrek. Najprej nekaj motivacije. Če je f merljiva in nenegativna, predpis

$$\nu(A) = \int_S f(x)\, 1_A(x)\, \mu(dx)$$

definira mero po izreku o monotoni konvergenci. Poleg tega velja, če je $\mu(A) = 0$, je tudi $\nu(A) = 0$. Kaj pa obratno?

<u>Definicija</u>: Mera ν je <u>absolutno zvezna</u> glede na mero μ, če iz $\mu(A) = 0$ sledi $\nu(A) = 0$.

Oznaka: $\nu \ll \mu$.

<u>Izrek 2.1</u> (Radon-Nikodým).

Naj bo μ σ-končna mera in $\nu \ll \mu$. Potem obstaja merljiva funkcija $f \geq 0$, da je

$$\nu(A) = \int_A f(x)\,\mu(dx)$$

Funkcija f je enolično določena do množic z mero 0 natančno. Poleg tega je $\int_S f(x)\,\mu(dx) < \infty$.

<u>Dokaz</u>: Odložimo.

<u>Izrek 2.2</u>: Naj bo $X \geq 0$ realna slučajna spremenljivka z $E(X) < \infty$ in $\mathcal{G} \subseteq \mathcal{F}$. Pogojna pričakovana vrednost obstaja.

<u>Dokaz</u>: Na σ-algebri $\mathcal{G}$ lahko definiramo dve meri: (i) mero P lahko zožimo na $\mathcal{G}$; (ii) definiramo $\nu(G) = E(X \cdot 1_G)$. Očitno je $\nu \ll P$ na $\mathcal{G}$. Izrek 2.1 zagotavlja obstoj $\mathcal{G}$-merljive funkcije $\tilde{X}$, za katero je

$$E(X \cdot 1_G) = E(\tilde{X} \cdot 1_G)$$

za vse G. Za splošen X z $E|X| < \infty$ trditev sledi, da vzamemo X^+ in X^-.

<u>Komentar</u>: Iz definicije takoj sledi tudi monotonost. Če je $X \leq Y$ za slučajni spremenljivki z $E|X| < \infty$ in $E|Y| < \infty$, je tudi

$$E(X|\mathcal{G}) \leq E(Y|\mathcal{G}) \quad \text{s.g.}$$

<u>Dokaz</u> Radon-Nikodýmovega izreka

Izrek velja za σ-končne mere, vendar se bomo omejili na končne mere.
Vzemimo najprej $\nu(A) \leq \mu(A)$ za vse $A \in \mathcal{F}$.
Naj bo $\{A_1, A_2, \ldots, A_n\}$ particija S.
Priredimo ji stopničasto funkcijo

$$p(x) = \begin{cases} \nu(A_k)/\mu(A_k), & x \in A_k \\ & \mu(A_k) > 0 \\ 0, & \text{sicer.} \end{cases}$$

Za vsako unijo A množic iz particije velja

$$\int_A p(x)\,\mu(dx) = \nu(A).$$

Naj bo $\{B_1, B_2, \ldots, B_m\}$ finejša particija od $\{A_1, \ldots, A_n\}$. in q pripadajoča stopničasta funkcija.

Ker je p na A_k konstantna, bo

$$\int_{A_k} p(x) \cdot q(x)\, \mu(dx) = \int_{A_k} p(x)^2\, \mu(dx)$$

in posledično

$$\int_S p(x)\, q(x)\, \mu(dx) = \int_S p^2(x)\, \mu(dx).$$

Sledi

$$\int_S q(x)^2\, \mu(dx) - \int_S p^2(x)\, \mu(dx)$$

$$= \int_S (q(x) - p(x))^2\, \mu(dx)$$

$$\geq 0$$

Definirajmo

$$\alpha := \sup\left\{ \int_S p^2(x)\, \mu(dx) \text{ po vseh particijah } S \right\}.$$

Ker za finejše particije dobimo večji integral, obstaja zaporedje P_n particij s pripadajočimi stopničastimi funkcijami p_n, da velja

$$\alpha \geq \int_S p_n(x)\,\mu(dx) \geq \alpha - 1/n .$$

Predpostavimo lahko $P_1 \subseteq P_2 \subseteq \ldots$.
To pomeni, da je za $n \geq m$

$$\int_S (p_m(x) - p_n(x))^2 \mu(dx)$$

$$= \int_S p_n^2(x)\,\mu(dx) - \int_S p_m^2(x)\,\mu(dx)$$

$$\leq \frac{1}{m}$$

Sledi, da je p_n Cauchyjevo zaporedje v $L^2(\mu)$. Torej obstaja funkcija f, da je $p_n \xrightarrow{L^2} f$.

Ker je $|p_n| \leq 1$, lahko privzamemo $0 \leq f \leq 1$. Ostane še dokaz, da je f prava funkcija. Naj bo A merljiva množica.

Vzemimo najbolj grobo finejšo partijo, ki vsebuje $\mathcal{P}_n$ in $\{A, A^c\}$. (tako lahko eksplicitno napišemo)

Naj bo q_n pripadajoča stopničasta funkcija. Velja

$$\nu(A) = \int_A q_n(x)\,\mu(dx)$$

$$= \int_A \left(q_n(x) - p_n(x) + p_n(x)\right)\mu(dx)$$

$$= \underbrace{\int_A \left(q_n(x) - p_n(x)\right)\mu(dx)}_{\to 0,\ \text{ko } n \to \infty} + \underbrace{\int_A p_n(x)\,\mu(dx)}$$

$$\to \int_A f(x)\,\mu(dx) \quad \text{ko } n \to \infty.$$

(utemelji bralec)

Odpraviti moramo še omejitev $\nu(A) \le \mu(A)$. Velja

$$\mu \le \mu + \nu \quad \text{in} \quad \nu \le \mu + \nu$$

Obstajata funkciji f_1 in f_2 z

$$\mu(A) = \int_A f_1(x)\,(\mu+\nu)(dx)$$

$$\nu(A) = \int_A f_2(x)\,(\mu+\nu)(dx).$$

Naj bo $N := \{x : f_1(x) = 0\}$.

Definirajmo

$$f(x) = \begin{cases} \dfrac{f_2(x)}{f_1(x)} & \text{za } x \in N^c \\ 0, & \text{sicer.} \end{cases}$$

Preverimo

$$\nu(A) = \nu(A \cap N^c) \quad (\text{zaradi } \nu \ll \mu)$$

$$= \int_{A \cap N^c} f_2(x)\,(\mu+\nu)(dx)$$

$$= \int_{A \cap N^c} f(x) \cdot f_1(x)\,(\mu+\nu)(dx)$$

$$= \int_{A \cap N^c} f(x)\,\mu(dx)$$

$$= \int_A f(x)\,\mu(dx).$$

Opomba: Na tiho smo uporabili, da velja

$$\int f(x)\,\nu(dx) = \int f(x)\cdot f_1(x)\,(\mu+\nu)(dx).$$

To zlahka preverimo.

Navedimo osnovne lastnosti pogojne pričakovane vrednosti.

Izrek 2.3: Naj bosta X_1, X_2 slučajni spremenljivki z $E|X_1| < \infty$ in $E|X_2| < \infty$. Velja:

$$E(\alpha X_1 + \beta X_2 \mid \mathcal{G})$$

$$= \alpha\, E(X_1 \mid \mathcal{G}) + \beta\, E(X_2 \mid \mathcal{G}).$$

Dokaz: Samo preverimo, da desna stran ustreza definiciji.

Izrek 2.4: Naj bo X slučajna spremenljivka z $E|X| < \infty$. Naj bo $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$. Velja

$$E[E(X|\mathcal{G})|\mathcal{H}]$$

$$= E(X|\mathcal{H})$$

Opomba: Tej lastnosti se reče stolpična lastnost.

Dokaz: Leva stran je $\mathcal{H}$-merljiva. Naj bo $H \in \mathcal{H}$ in računamo

$$E[E[E(X|\mathcal{G})|\mathcal{H}] \cdot 1_H]$$

$$= E[E(X|\mathcal{G}) \cdot 1_H]$$

$$= E[X \cdot 1_H].$$

Izrek 2.5: Naj bo X slučajna spremenljivka z $E|X| < \infty$. Naj bo U $\mathcal{G}$-merljiva in naj velja $E|U \cdot X| < \infty$. Velja

$$E(U \cdot X \mid \mathcal{G}) = U \cdot E(X \mid \mathcal{G}).$$

Dokaz: Za stopničaste U je trditev očitna. Obstaja zaporedje stopničastih U_n z $|U_n| \leq |U|$ in $U_n \to U$, kjer so U_n $\mathcal{G}$-merljive. Po izreku o dominirani konvergenci bo veljalo

$$\begin{array}{ccc} E(U_n \cdot X \cdot 1_G) & \to & E(U \cdot X \cdot 1_G) \\ \parallel & & \\ E[U_n E(X \mid \mathcal{G}) 1_G] & \to & E(U E(X \mid \mathcal{G}) 1_G) \end{array}$$

Trditev sledi, če pokažemo, da je $E[|U \cdot E(X \mid \mathcal{G})|] < \infty$.

Po Fatouju je

$$E[\,|u \cdot E(X|\mathcal{G})|\,]$$

$$= E[\liminf_{n\to\infty} |u_n E(X|\mathcal{G})|]$$

$$\leq \liminf_{n\to\infty} E[\,|u_n E(X|\mathcal{G})|\,]$$

$$= \liminf_{n\to\infty} E[\,|E(X \cdot u_n)|\,]$$

$$\leq \liminf_{n\to\infty} E|X \cdot u_n|$$

$$\leq \liminf_{n\to\infty} E|X \cdot u| < \infty.$$

Zaključne pripombe:

(i) Če je $\mathcal{G} = \sigma(X)$, obstaja merljiva funkcija ψ, da je $E(Y|\mathcal{G}) = \psi(X)$. Isto velja za $\sigma(X_1, X_2, \dots, X_n)$.

(ii) Če je $\mathcal{G} = \sigma(\mathcal{P})$, kjer je $\mathcal{P}$ π-sistem, je dovolj preveriti

$$E(X \cdot 1_G) = E(E(X|\mathcal{G})1_G)$$

za $G \in \mathcal{P}$.

<u>Primer</u>: σ-algebro $\sigma(X, Y)$ generira π-sistem

$$\mathcal{P} = \{G \cap H : G \in \sigma(X),\ H \in \sigma(Y)\}.$$

2.2. Pogojne porazdelitve

<u>Motivacija</u>: za diskretni X in Y smo pogojno porazdelitev Y glede na $\{X = x\}$ definirali z

$$P(Y=y \mid X=x) = \frac{P(X=x, Y=y)}{P(X=x)} .$$

Za vsak $A \in \mathcal{B}(\mathbb{R})$ velja

$$E(1_A(Y) \mid X=x) = \sum_{y \in A} P(Y=y \mid X=x)$$

$$= P(Y \in A \mid X=x)$$

$$= \varphi_A(x)$$

Če zamenjamo X z diskretno spr. X', za katero je $\sigma(X) = \sigma(X')$, bo

$$\varphi_A(X) = \varphi'_A(X') .$$

<u>Ideja</u>: Pogojna porazdelitev bo mera „odvisna" samo od $\sigma(X)$. Ta mera bo odvisna od $\omega \in \Omega$.

Zgoraj smo videli, da bi moralo veljati

$$E(1_A(Y) \mid \mathcal{G}) = \int 1_A(y)\, Q(dy, \omega),$$

če je Q mera odvisna od ω. Vendar je desna stran kar $Q(A, \omega)$, ki mora biti $\mathcal{G}$-merljiva. To nas napelje na naslednjo definicijo.

<u>Definicija</u>: Naj bo Y slučajna spremenljivka na $(\Omega, \mathcal{F}, P)$ z vrednostmi v $(S, \mathcal{S})$. Pogojna porazdelitev Y glede na $\mathcal{G} \subseteq \mathcal{F}$ je jedro (preslikava)

$$Q : \mathcal{S} \times \Omega \to [0,1],$$

da velja:

(i) za skoraj vsak $\omega \in \Omega$ je $Q(\cdot, \omega)$ verjetnostna mera.

(ii) za $A \in \mathcal{S}$ je

$$Q(A, \cdot) = E[1_A(Y) \mid \mathcal{G}].$$

Opombe:

(i) desna stran v (ii) vedno obstaja za fiksen A. Vendar to ni dovolj za obstoj Q, ker je množic A „preveč".

(ii) Če je f omejena merljiva funkcija, sledi

$$E(f(X) \mid \mathcal{G}) = \int_S f(y)\, Q(dy, \cdot)$$

Za stopničaste funkcije to velja po definiciji. Sicer pa preverimo z aproksimacijami.

Kako je z obstojem Q? Izkaže se, da Q ne obstaja vedno. Prostor $(S, \mathcal{S})$ mora biti dovolj „lep".

Izrek 2.6: Naj bo $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ in $Y : \Omega \to \mathbb{R}$. Naj bo $\mathcal{G} \subseteq \mathcal{F}$. Pogojna porazdelitev Q obstaja.

Dokaz: Za vsako racionalno število $r \in \mathbb{Q}$ obstaja $E(1_{(-\infty, r]}(Y) \mid \mathcal{G})$. Zaradi monotonosti je za $s < r$, $s, r \in \mathbb{Q}$

$$E[1_{(-\infty, s]}(Y) \mid \mathcal{G}] \leq E[1_{(-\infty, r]}(Y) \mid \mathcal{G}]$$

s.g. Tovej lahko vse te pogojne pričakovane vrednosti popravimo, da bo razen na množici N z $P(N) = 0$ veljalo, da je funkcija

$$q \longmapsto E(1_{(-\infty, q]}(Y) \mid \mathcal{G}) = G(q, \omega)$$

nepadajoča. Definiramo za $x \in \mathbb{R}$

$$F(x, \omega) = \inf\{G(q) : q > x\}.$$

Funkcija F je nepadajoča in desno zvezna. „Uredimo" lahko še to, da je $F(x, \omega) \to 1$, ko $x \to \infty$ in $F(x, \omega) \to 0$, ko $x \to -\infty$ s.g.

F je porazdelitvena funkcija, ki ji enolično pripada mera $Q(\cdot, \omega)$. Pokazati moramo, da je Q res pogojna porazdelitev. Pokazati moramo dvoje:

(i) $F(x, \omega)$ je verzija

$$E[1_{(-\infty, x]}(Y) \mid \mathcal{G}]$$

za vse $x \in \mathbb{R}$.

(ii) Družina $\mathcal{A} = \{A \in \mathcal{B}(\mathbb{R}) :$
$Q(A, \omega)$ je $\mathcal{G}$-merljiva in
$Q(A, \omega)$ je verzija
$E(1_A(Y) \mid \mathcal{G})\}$

vsebuje $\mathcal{B}(\mathbb{R})$.

Poglejmo najprej trditev (ii).

Po definiciji $\mathcal{A}$ vsebuje končne disjunktne unije $\bigcup_{i=1}^{n} (a_i, b_i]$.

Z izrekom o monotoni konvergenci preverimo, da je

$\mathcal{A}$ λ-sistem. To zaključi dokaz.

Pospraviti moramo še (i). Naj bo $q_n \in \mathbb{Q}$ in $q_n \downarrow x$. Potem

$$G(q_n, \omega) \downarrow F(x, \omega), \text{ ko } n \to \infty.$$

Zaradi monotonosti s.g.

$$E[1_{(-\infty, q_n]}(Y) \mid \mathcal{G}] \downarrow E[1_{(-\infty, x]}(Y) \mid \mathcal{G}]$$

ko $n \to \infty$.

<u>Opomba</u>: Uporabili smo dejstvo, da za omejene X_n z $X_n \to X$ velja

$$E(X_n \mid \mathcal{G}) \to E(X \mid \mathcal{G}) \text{ s.g.}$$

To pomeni, da je $F(x, \omega)$ verzija $E(1_{(-\infty, x]}(Y) \mid \mathcal{G})$.

Dodatek: Jensenova neenakost.

Če je $P(X \in (a,b)) = 1$ in je

$\varphi : (a,b) \to \mathbb{R}$ konveksna funkcija

z $E|\varphi(X)| < \infty$, velja za $E|X| < \infty$

$$\varphi(E(X)) \leq E(\varphi(X)).$$

Neenačba velja tudi v pogojni verziji, vendar z dostavkom s.g.

Velja namreč

$$E(\varphi(X) \mid \mathcal{G})(\omega) = \int_{\mathbb{R}} \varphi(x)\, Q(dx, \omega) \quad \text{s.g.}$$

Integral na desni obstaja po absolutni vrednosti s.g (zakaj?).

Po navadni Jensenovi neenačbi je

$$\int_{\mathbb{R}} \varphi(x)\, Q(dx, \omega) \geq \varphi\left(\int_{\mathbb{R}} x\, Q(dx, \omega)\right)$$

$$= \varphi(E(X \mid \mathcal{G})(\omega)).$$