Rachunek prawdopodobieństwa MAT1332

Wydział Matematyki, Matematyka Stosowana

Wykładowca: dr hab. Agnieszka Jurlewicz

Wykład 12: Warunkowa wartość oczekiwana. Rozkłady warunkowe. Mieszanina rozkładów.

Warunkowa wartość oczekiwana.

Definicja:

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną. **Warunkową wartością oczekiwaną** (w.w.o.) zmiennej losowej X względem σ-ciała $\mathcal{G} \subset \mathcal{F}$ nazywamy zmienną losową $\mathrm{E}(X|\mathcal{G})$, która jest $\mathcal{G}$-mierzalna oraz spełnia warunek

$$\int_G \mathrm{E}(X|\mathcal{G})\, dP = \int_G X\, dP \ \ \forall G \in \mathcal{G}.$$

Twierdzenie: Jeżeli $\mathrm{E}|X| < \infty$, to warunkowa wartość oczekiwana $\mathrm{E}(X|\mathcal{G})$ jest dobrze określona z tw. Radona-Nikodyma.

Własności warunkowej wartości oczekiwanej:

Przy założeniu, że $\mathrm{E}|X| < \infty$

(a) $\mathrm{E}(\mathrm{E}(X|\mathcal{G})) = \mathrm{E}X$;

(b) jeżeli $X \geqslant 0$ p.n., to $\mathrm{E}(X|\mathcal{G}) \geqslant 0$ p.n.;

(c) $\mathrm{E}(aX + bY|\mathcal{G}) = a\mathrm{E}(X|\mathcal{G}) + b\mathrm{E}(Y|\mathcal{G})$ p.n., gdzie a, b są dowolnymi stałymi.

(d) jeżeli X jest $\mathcal{G}$-mierzalna, to $\mathrm{E}(X|\mathcal{G}) = X$ p.n.;

(e) jeżeli X jest niezależna od $\mathcal{G}$, to $\mathrm{E}(X|\mathcal{G}) = \mathrm{E}X$ p.n.

(f) dla $\mathcal{G}_1 \subset \mathcal{G} \subset \mathcal{F}$ mamy $\mathrm{E}(\mathrm{E}(X|\mathcal{G}_1)|\mathcal{G}) = \mathrm{E}(\mathrm{E}(X|\mathcal{G})|\mathcal{G}_1) = \mathrm{E}(X|\mathcal{G}_1)$ p.n.

(g) $\mathrm{E}|\mathrm{E}(X|\mathcal{G})| \leqslant \mathrm{E}|X|$;

(h) jeżeli $X_n \xrightarrow{L^1} X$, to $\mathrm{E}(X_n|\mathcal{G}) \xrightarrow{L^1} \mathrm{E}(X|\mathcal{G})$.

(i) jeżeli $X_n \to X$ p.n., $|X_n| \leqslant Y$ dla pewnej $Y \in L^1(P)$,
to $\mathrm{E}(X_n|\mathcal{G}) \to \mathrm{E}(X|\mathcal{G})$ p.n.;

(j) jeżeli $X_n \geqslant 0$ tworzą ciąg niemalejący, $X_n \to X$ p.n. dla $X \in L^1(P)$,
to $\mathrm{E}(X_n|\mathcal{G}) \to \mathrm{E}(X|\mathcal{G})$ p.n.

(k) jeżeli X jest $\mathcal{G}$-mierzalna, $\mathrm{E}|XY| < \infty$, to $\mathrm{E}(XY|\mathcal{G}) = X\mathrm{E}(Y|\mathcal{G})$ p.n.;

(l) jeżeli X jest $\mathcal{G}$-mierzalna, $g(x, y)$ jest funkcją borelowską, $\mathrm{E}|g(X, Y)| < \infty$,
to $\mathrm{E}(g(X, Y)|\mathcal{G}) = \mathrm{E}(g(x, Y)|\mathcal{G})|_{x=X}$ p.n.

Opracowanie: dr hab. Agnieszka Jurlewicz

Niech Y będzie zmienną losową, $\sigma(Y) = \sigma\{Y^{-1}(B), B \in \mathcal{B}_R\}$, gdzie $\mathcal{B}_R$ to σ-ciało zbiorów borelowskich. Wówczas $\mathrm{E}(X|\sigma(Y))$ oznaczamy krótko $\mathrm{E}(X|Y)$.

Fakt: Istnieje funkcja borelowska m, taka że $\mathrm{E}(X|Y) = m(Y)$ p.n.;

Uwaga:

Często podajemy warunkową wartość oczekiwana tylko poprzez wzór na funkcję $m(y)$, przy czym zapis ma postać:

$$\mathrm{E}(X|Y = y) = m(y).$$

Funkcja $m(y)$ zwana jest **funkcją regresji I rodzaju**.

Fakt: Jeżeli $\mathrm{D}^2Y < \infty$, $\mathrm{D}^2X < \infty$, to $\min\limits_f \mathrm{E}(X - f(Y))^2 = \mathrm{E}(X - m(Y))^2$
(dla f - funkcji borelowskich).

Rozkłady warunkowe.

Definicja: Niech B będzie zbiorem borelowskim, a X pewną zmienną losową. Definiujemy nową zmienną losową $Z = \mathbb{1}_{\{X \in B\}}$.

Rozkładem warunkowym zmiennej losowej X względem σ-ciała $\mathcal{G}$ nazywamy $P(X \in B|\mathcal{G}) := \mathrm{E}(Z|\mathcal{G}) = \mathrm{E}(\mathbb{1}_{\{X\in B\}}|\mathcal{G})$, jako funkcję borelowskiego zbioru B.

- Zauważmy, że $\mathrm{E}|Z| = \mathrm{E}Z = P(X \in B) < \infty$ dla dowolnego B, zatem warunkowa wartość oczekiwana $\mathrm{E}(Z|\mathcal{G})$ (czyli rozkład warunkowy Y względem $\mathcal{G}$) jest dobrze zdefiniowana.

- Dla $\mathcal{G} = \sigma(Y)$, gdzie Y jest pewną zmienną losową, $P(X \in B|\mathcal{G}) = P(X \in B|Y)$ nazywamy rozkładem warunkowym zmiennej losowej X pod warunkiem Y.

- Przy ustalonym y (takim, który jest wartością istotnie przyjmowaną przez Y) rozkład warunkowy $P(X \in B|Y = y)$ jako funkcja zbioru borelowskiego B jest pewnym rozkładem prawdopodobieństwa. Rozkład ten ma dystrybuantę

  $$F(x|y) = P(X < x|Y = y)$$

  (zwaną **dystrybuantą warunkową**).

  Rozkład warunkowy X pod warunkiem Y możemy zatem opisać podając **rodzinę dystrybuant warunkowych** $F(x|y)$ po wszystkich wartościach y istotnie przyjmowanych przez Y.

- Dla dowolnej funkcji borelowskiej h takiej, że $\mathrm{E}|h(X)| < \infty$ mamy

  $$\mathrm{E}(h(X)|Y = y) = \int\limits_{\mathbb{R}} h(x)\, F(dx|y).$$

  W szczególności, warunkowa wartość oczekiwana $\mathrm{E}(X|Y = y)$ przy ustalonym y to zwykła wartość oczekiwana rozkładu warunkowego $P(X \in B|Y = y)$.

Opracowanie: dr hab. Agnieszka Jurlewicz

- Jeżeli wektor losowy (X, Y) ma rozkład **dyskretny** zadany ciągiem $\{(x_n, y_k, p_{nk}), n \in \mathbb{T}_1, k \in \mathbb{T}_2\}$, to rozkład warunkowy X pod warunkiem $Y = y_k$ też jest dyskretny i opisać możemy go także podając **rodzinę ciągów**

$$\left\{ \left(\left(x_n, \frac{p_{nk}}{p_{\cdot k}} \right) \middle| y_k \right), n \in \mathbb{T}_1 \right\}$$

po wszystkich tych wartościach y_k, dla których $p_{\cdot k} = P(Y = y_k) > 0$, czyli po wartościach istotnie przyjmowanych przez Y.

(Zauważmy, że $\frac{p_{nk}}{p_{\cdot k}} = P(X = x_n | Y = y_k)$ to zwykłe prawdopodobieństwo warunkowe.)

- Jeżeli wektor losowy (X, Y) ma rozkład **ciągły** o gęstości $f_{X,Y}(x, y)$, to rozkład warunkowy X pod warunkiem $Y = y$ też jest ciągły o gęstości

$$f(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

dla wszystkich takich y, dla których $f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)dx > 0$, czyli po wartościach istotnie przyjmowanych przez Y.

Rozkład warunkowy opisujemy zatem wtedy także podając **rodzinę gęstości warunkowych** $f(x|y)$ po wszystkich wartościach y, dla których $f_Y(y) > 0$.

- Wzór na prawdopodobieństwo całkowite:

$$P(X \in B) = \int\limits_{-\infty}^{\infty} P(X \in B | Y = y) dF_Y(y).$$

gdzie F_Y to dystrybuanta rozkładu zmiennej losowej Y.

- Jeśli znamy rozkład zmiennej losowej Y i rozkład warunkowy X pod warunkiem Y, to znamy też rozkład łączny wektora losowego (X, Y):

$$F_{X,Y}(x, y) = \int\limits_{-\infty}^{y} F(x|y') dF_Y(y'). \tag{1}$$

Dla rozkładu dyskretnego $p_{nk} = P(X = x_n | Y = y_k) p_{\cdot k}$;
dla rozkładu ciągłego $f_{X,Y}(x, y) = f(x|y) f_Y(y)$.

- Jeżeli X i Y są **niezależnymi** zmiennymi losowymi, to $P(X \in B|Y) = P(X \in B)$ z prawd. 1, tzn. wtedy rozkład warunkowy jest taki sam jak rozkład zmiennej losowej X. Wtedy wzór (1) sprowadza się do znanego wzoru $F_{X,Y}(x, y) = F_X(x) F_Y(y)$.

Opracowanie: dr hab. Agnieszka Jurlewicz

SUMA LOSOWA.

Niech $X_1,\ X_2,\ldots$ będzie ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie. Niech N będzie indeksem losowym (tzn. zmienną losową dyskretną przyjmującą tylko wartości naturalne $1,2,\ldots$) niezależną od ciągu $\{X_k, k=1,2,\ldots\}$.
Sumą losową nazywamy zmienną losową

$$S=\sum_{k=1}^{N} X_k = X_1+X_2+\ldots+X_N.$$

(N to losowa ilość składników w sumie.)

- Załóżmy, że istnieją skończone wartości oczekiwane $\mathrm{E}X_1$ i $\mathrm{E}N$. Korzystając z techniki warunkowej wartości oczekiwanej obliczymy wartość oczekiwaną $\mathrm{E}S$:

$$\mathrm{E}S=\mathrm{E}(\mathrm{E}(S|N))=\mathrm{E}\left(\mathrm{E}\left(\sum_{k=1}^{N} X_k\Big|N\right)\right)=\mathrm{E}\left(\mathrm{E}\left(\sum_{k=1}^{n} X_k\Big|N\right)\Bigg|_{n=N}\right)=$$
$$=\mathrm{E}\left(\mathrm{E}\left(\sum_{k=1}^{n} X_k\right)\Bigg|_{n=N}\right)=\mathrm{E}\left((n\mathrm{E}X_1)\Bigg|_{n=N}\right)=\mathrm{E}(N\cdot\mathrm{E}X_1)=\mathrm{E}N\cdot\mathrm{E}X_1.$$

Otrzymaliśmy zatem, że
$$\mathrm{E}S=\mathrm{E}N\cdot\mathrm{E}X_1.$$

- Załóżmy teraz, że $\mathrm{D}^2X_1<\infty$ i $\mathrm{D}^2N<\infty$. Obliczymy wariancję D^2S:

$$\mathrm{E}S^2=\mathrm{E}(\mathrm{E}(S^2|N))=\mathrm{E}\left(\mathrm{E}\left(\left(\sum_{k=1}^{N} X_k\right)^2\Big|N\right)\right)=\mathrm{E}\left(\mathrm{E}\left(\left(\sum_{k=1}^{n} X_k\right)^2\Big|N\right)\Bigg|_{n=N}\right)=$$
$$=\mathrm{E}\left(\mathrm{E}\left(\sum_{k=1}^{n} X_k\right)^2\Bigg|_{n=N}\right)=\mathrm{E}\left(\mathrm{E}\left(\sum_{k=1}^{n} X_k^2+\sum_{k=1}^{n}\sum_{\substack{j=1\\ j\neq k}}^{n} X_kX_j\right)\Bigg|_{n=N}\right)=$$
$$=\mathrm{E}\left((n\mathrm{E}X_1^2+n(n-1)(\mathrm{E}X_1)^2)\Bigg|_{n=N}\right)=\mathrm{E}(N\cdot\mathrm{E}X_1^2+N(N-1)(\mathrm{E}X_1)^2)=$$
$$=\mathrm{E}N\cdot(\mathrm{E}X_1^2-(\mathrm{E}X_1)^2)+\mathrm{E}N^2\cdot(\mathrm{E}X_1)^2=\mathrm{E}N\cdot\mathrm{D}^2X_1+\mathrm{E}N^2\cdot(\mathrm{E}X_1)^2.$$

Zatem
$$\mathrm{D}^2S=\mathrm{E}N\cdot\mathrm{D}^2X_1+\mathrm{E}N^2\cdot(\mathrm{E}X_1)^2-(\mathrm{E}N\cdot\mathrm{E}X_1)^2=$$
$$=\mathrm{E}N\cdot\mathrm{D}^2X_1+(\mathrm{E}N^2-(\mathrm{E}N)^2)\cdot(\mathrm{E}X_1)^2=\mathrm{E}N\cdot\mathrm{D}^2X_1+\mathrm{D}^2N\cdot(\mathrm{E}X_1)^2.$$

Otrzymaliśmy zatem, że

$$\mathrm{D}^2S=\mathrm{E}N\cdot\mathrm{D}^2X_1+\mathrm{D}^2N\cdot(\mathrm{E}X_1)^2.$$

- Możemy też określić rozkład S za pomocą funkcji charakterystycznej $\varphi_S(t)$.

  Niech $\varphi_X(t)$ oznacza funkcję charakterystyczną zmiennych losowych X_k,
  a $g_N(z)$ - funkcję tworzącą losowego indeksu N. Wtedy

$$\varphi_S(t)=g_N(\varphi_X(t))$$

Opracowanie: dr hab. Agnieszka Jurlewicz

UZASADNIENIE:

$$\varphi_S(t) = \mathrm{E}e^{itS} = \mathrm{E}(\mathrm{E}(e^{itS}|N)) = \mathrm{E}\left(\mathrm{E}\left(\exp\left(it\sum_{k=1}^{n} X_k\right)\Big|N\right)\Big|_{n=N}\right) =$$
$$= \mathrm{E}\left(\mathrm{E}\exp\left(it\sum_{k=1}^{n} X_k\right)\Big|_{n=N}\right) = \mathrm{E}\left(\varphi_{X_1+\ldots+X_n}(t)\Big|_{n=N}\right) = \mathrm{E}\left((\varphi_X(t))^n\Big|_{n=N}\right) =$$
$$= \mathrm{E}\left(\varphi_X(t)\right)^N = g_N(\varphi_X(t))$$

MIESZANINA ROZKŁADÓW.

DEFINICJA:

Niech N będzie zmienną losową dyskretną przyjmującą wartości naturalne (**indeksem losowym**), przy czym $p_n = P(N=n)$ dla $n = 1, 2, \ldots, n_0$, $\sum\limits_{n=1}^{n_0} p_n = 1$.

Y_1, $Y_2, \ldots$, Y_{n_0} to ciąg zmiennych losowych niezależnych od N, o dystrybuantach odpowiednio F_1, $F_2, \ldots$, F_{n_0}.

Rozkład zmiennej losowej $Z = Y_N$ nazywamy **mieszaniną rozkładów** F_1, $F_2, \ldots$, F_{n_0}, przy czym p_1, $p_2, \ldots$, p_{n_0} to **udziały tych rozkładów w mieszaninie**.

UWAGA:

Mamy $P(Z = Y_n) = P(N = n) = p_n$ oraz

$F_Z(z) = P(Z < z) = \sum\limits_{n=1}^{n_0} P(Z < z, N = n) = \sum\limits_{n=1}^{n_0} P(Y_n < z, N = n) =$

$=$[z niezależności N i Y_n] $= \sum\limits_{n=1}^{n_0} P(Y_n < z)P(N = n) = \sum\limits_{n=1}^{n_0} p_n F_n(z)$, czyli

$$F_Z(z) = p_1 F_1(z) + \ldots + p_{n_0} F_{n_0}(z).$$

Natomiast $\mathrm{E}Z = \mathrm{E}(\mathrm{E}(Z|N)) = \mathrm{E}((\mathrm{E}Y_n)|_{n=N}) = \sum\limits_{n=1}^{n_0} p_n \mathrm{E}Y_n$, czyli

$$\mathrm{E}Z = p_1\mathrm{E}Y_1 + \ldots + p_{n_0}\mathrm{E}Y_{n_0}.$$

ROZKŁAD MIESZANY

TWIERDZENIE: Każdy rozkład prawdopodobieństwa jest mieszaniną rozkładu dyskretnego, rozkładu ciągłego i rozkładu osobliwego, tzn. dowolną dystrybuantę $F(x)$ można w jednoznaczny sposób przedstawić jako

$$F(x) = p_d F_d(x) + p_c F_c(x) + p_o F_o(x)$$

dla pewnych stałych p_d, p_c, $p_o \geqslant 0$ takich, że $p_d + p_c + p_o = 1$, oraz pewnych dystrybuant F_d, F_c, F_o odpowiednio rozkładu dyskretnego, ciągłego, osobliwego.

DEFINICJA: **Rozkład mieszany** to taki, dla którego przynajmniej dwie spośród liczb p_d, p_c, p_o są większe od 0.

Opracowanie: dr hab. Agnieszka Jurlewicz