Κατανομή ακραίων τιμών Gumbel -για ακραίες τιμές

Κείμενα από Τσακίρης, 2013 και Μπέλλος, 2010

2.6.2 Κατανομή Ακραίων Τιμών Τύπου I (Gumbel)

Συνάρτηση πυκνότητας πιθανότητας (p.d.f.):

$$p(x) = \exp\left\{ \mp \frac{x-\beta}{\alpha} - \exp\left[\mp \frac{x-\beta}{\alpha} \right] \right\} \tag{2.21}$$

$$-\infty < x < +\infty, \quad -\infty < \beta < +\infty, \quad \alpha > 0$$

όπου α είναι η παράμετρος κλίμακας και β είναι παράμετρος θέσεως. Τα σύμβολα + και − είναι: (−) για μέγιστα και (+) για ελάχιστα.

Βασικές στατιστικές ποσότητες:

Προσδοκώμενη τιμή $E(X) = \beta + 0.577\,\alpha$ (μέγιστα)

(μέσος όρος) $E(X) = \beta - 0.577\,\alpha$ (ελάχιστα)

Διασπορά $\mathrm{Var}(X) = 1.645\,\alpha^2$ (μέγιστα και ελάχιστα) (2.22)

Συντελεστής ασυμμετρίας $g = 1.1396$ (μέγιστα)

$g = -1.1396$ (ελάχιστα)

Με την αντικατάσταση:

$$y = \frac{x - \beta}{\alpha} \tag{2.23}$$

p.d.f. $\quad p(y) = \exp\left[\mp y - \exp(\mp y)\right]$ (2.24)

c.d.f. $\quad P(y) = \int_{-\infty}^{y} \exp\left[\mp y - \exp(\mp y)\right] dy$ (2.25)

Τελικά

$$P(y) = \exp(-\exp(-y)) \quad \text{(μέγιστα)} \tag{2.26}$$

$$P(y) = 1 - \exp(-\exp y) \quad \text{(ελάχιστα)} \tag{2.27}$$

Με βάση τις τελευταίες εξισώσεις έχει γίνει η πινακοποίηση της αθροιστικής πιθανότητας (U.S. National Bureau of Standards, 1953).

Με τη μέθοδο των ροπών η εκτίμηση των παραμέτρων γίνεται ως εξής:

$$\widehat{\alpha} = \frac{\widehat{\sigma}}{1.283} \quad \text{και} \tag{2.28}$$

Μέθοδος Ροπών/ εκτίμηση παραμέτρων

Με τη μέθοδο των ροπών η εκτίμηση των παραμέτρων γίνεται ως εξής:

$$\hat{\alpha} = \frac{\hat{\sigma}}{1.283} \quad \text{και} \tag{2.28}$$

$$\hat{\beta} = \bar{x} - 0.45\,\hat{\sigma} \quad \text{(μέγιστα)} \tag{2.29}$$

$$\hat{\beta} = \bar{x} + 0.45\,\hat{\sigma} \quad \text{(ελάχιστα)} \tag{2.30}$$

Παράδειγμα. Σειρά ετησίων μεγίστων υψών βροχής μιας ορισμένης διάρκειας έχει μέσο όρο $\bar{x} = 14.19$ mm και τυπική απόκλιση $\hat{\sigma} = 8.46$ mm. Χρησιμοποιώντας την κατανομή Gumbel να προσδιορισθεί το ύψος που αντιστοιχεί σε περίοδο επαναφοράς $T = 10$ έτη. Σύμφωνα με τη μέθοδο των ροπών $\hat{\alpha} = \hat{\sigma}/1.283 = 6.594$, $\hat{\beta} = \bar{x} - 0.45\,\hat{\sigma} = 10.38$. Επίσης $P(y) = 1 - (1/T) = 0.90$ και επομένως $\exp[-\exp(-y)] = 0.90$ ή $y = -\ln[-\ln 0.90] = 2.25$. Τέλος $x = \hat{\alpha}\,y + \hat{\beta} = 6.594 \cdot 2.25 + 10.38 = 25.22$ mm.

Μέθοδος Παράγοντα Συχνότητας

Με τη μέθοδο του παράγοντα συχνότητας για πολύ μεγάλο δείγμα ($N \to \infty$) έχει προκύψει αναλυτικά η σχέση μεταξύ παράγοντα συχνότητας και περιόδου επαναφοράς

$$K_T = -0.7797\left[0.5772 + \ln\left(\ln\left[T/(T-1)\right]\right)\right] \tag{2.31}$$

Για μικρά σχετικά δείγματα ($N < 100$) ο παράγοντας συχνότητας υπολογίζεται ως εξής

$$K_T = \frac{-\left[\ln\left(\ln\left[T/(T-1)\right]\right) + \overline{y}_N\right]}{\sigma_N} \tag{2.32}$$

όπου $\overline{y}_N$ και σ_N (μέσος όρος και τυπική απόκλιση της ανηγμένης μεταβλητής y) βρίσκονται στον Πίν. 2.3 με βάση τον αριθμό των παρατηρήσεων του δείγματος.

Αν στο προηγούμενο παράδειγμα εφαρμόσουμε τη μεθοδολογία του Παράγοντα Συχνότητας (Εξ. 2.31 και 2.32) τότε προκύπτουν τα εξής αποτελέσματα:

Από την Εξ. 2.31

$$K_T = -0.7797\left[0.5772 + \ln\left(\ln\frac{10}{9}\right)\right] = 1.3046$$

Επίσης

$$c_v = \frac{\widehat{\sigma}}{\overline{x}} = \frac{8.46}{14.19} = 0.5962$$

Συνεπώς σύμφωνα με την Εξ. 2.16

$$x_{10} = 14.19(1 + 0.5962 \cdot 1.3046) = 25.23 \text{ mm}$$

$X_T = \tilde{\mu} + \tilde{\sigma} K_T$ – Μέθοδος παράγοντα συχνότητας

Εφαρμογή

Άσκηση 2.3

Δίνονται τα ετήσια μέγιστα των μέσων ημερήσιων παροχών ενός υδατορεύματος για την περίοδο 1930-1970.

Πίν. 2.10: *Ετήσια μέγιστα μέσων ημερήσιων παροχών*

α/α	Έτος	Μέγιστη παροχή (m^3/s)	α/α	Έτος	Μέγιστη παροχή (m^3/s)	α/α	Έτος	Μέγιστη παροχή (m^3/s)
1	1930	67	15	1944	102	29	1958	66
2	1931	123	16	1945	60	30	1959	22
3	1932	145	17	1946	32	31	1960	37
4	1933	55	18	1947	109	32	1961	39
5	1934	21	19	1948	40	33	1962	43
6	1935	47	20	1949	68	34	1963	49
7	1936	73	21	1950	42	35	1964	21
8	1937	103	22	1951	66	36	1965	38
9	1938	86	23	1952	84	37	1966	21
10	1939	49	24	1953	45	38	1967	32
11	1940	75	25	1954	38	39	1968	74
12	1941	46	26	1955	51	40	1969	75
13	1942	39	27	1956	57	41	1970	24
14	1943	36	28	1957	200			

Λύση

α) Υπολογίζουμε το μέσο όρο και την τυπική απόκλιση του δείγματος:

Μέσος όρος:

$$\overline{Q} = \sum_{i=1}^{N} \frac{Q_i}{N}$$

Τυπική απόκλιση:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{N} (Q_i - \overline{Q})^2}{N-1}}$$

όπου ο μέσος όρος των ετήσιων μεγίστων των μέσων ημερήσιων παροχών του υδατορεύματος για την περίοδο 1930-1970, $\hat{\sigma}$ η τυπική απόκλιση του δείγματος και N το πλήθος των τιμών του δείγματος. Το πλήθος των τιμών του δείγματος είναι $N = 41$.

Επομένως ο μέσος όρος και η τυπική απόκλιση του συγκεκριμένου δείγματος είναι:

$$\overline{Q} = 61 \text{ m}^3/\text{s}$$

$$\hat{\sigma} = 36 \text{ m}^3/\text{s}$$

Από το δείγμα
Μέσος όρος τυπική απόκλιση

Μέθοδος ροπών, εκτίμηση παραμέτρων

i. Μέθοδος ροπών

Η εκτίμηση των παραμέτρων για τη μέθοδο των ροπών γίνεται με βάση τους τύπους:

$$\hat{a} = \frac{\hat{\sigma}}{1.283}$$

$$\hat{\beta} = \bar{x} - 0.45 \cdot \hat{\sigma} \quad \text{για μέγιστα}$$

ενώ η αθροιστική πιθανότητα υπολογίζεται σύμφωνα με την εξίσωση:

$$P(y) = \exp(-\exp(-y)) \quad \text{(για μέγιστα)}$$

Έχουμε:

$$\hat{a} = \frac{\hat{\sigma}}{1.283} \Rightarrow \hat{a} = \frac{36}{1.283} \Rightarrow \hat{a} = 28.059$$

$$\hat{\beta} = \bar{x} - 0.45 \cdot \hat{\sigma} \Rightarrow \hat{\beta} = 61 - 0.45 \cdot 36 \Rightarrow \hat{\beta} = 44.800$$

Για την περίοδο επαναφοράς $T = 65$ έτη, η αθροιστική πιθανότητα είναι:

$$P(y) = P(Q \leq q) = 1 - \frac{1}{T} \Rightarrow P(y) = 1 - \frac{1}{65} \Rightarrow P(y) = 0.985$$

Αλλά από τις εξισώσεις της κατανομής έχουμε:

$$P(y) = \exp(-\exp(-y)) \Rightarrow \exp(-\exp(-y)) = 0.985 \Rightarrow$$
$$\Rightarrow y = \ln(-\ln(0.985)) \Rightarrow y = 4.192.$$

Επίσης

$$y = \frac{x - \widehat{\beta}}{\hat{a}} \Rightarrow y = \frac{Q_{65} - \widehat{\beta}}{\hat{a}} \Rightarrow Q_{65} = y \cdot \hat{a} + \widehat{\beta}$$

Η παροχή με περίοδο επαναφοράς 65 έτη είναι τελικά:

$$Q_{65} = y \cdot \hat{a} + \widehat{\beta} \Rightarrow Q_{65} = 4.192 \cdot 28.059 + 44.800 \Rightarrow$$
$$\Rightarrow Q_{65} = 162.4 \text{ m}^3/\text{s}.$$

Από πιθανότητα----
μέγεθος (αντίστροφο
πρόβλημα), εύρεση y
από αθροιστική
συνάρτηση
πιθανότητας

Ο παράγοντας συχνότητας k_T για μικρά σχετικά δείγματα ($N < 100$), υπολογίζεται από τη σχέση:

$$k_T = \frac{-\left[\ln\left(\ln\frac{T}{T-1}\right) + \overline{y}_N\right]}{\sigma_N}$$

όπου $\overline{y}_N$ και σ_N (μέσος όρος και τυπική απόκλιση της ανηγμένης μεταβλητής y), που δίνονται από τον Πίνακα 2.2 με βάση των αριθμό των παρατηρήσεων του δείγματος:

$$\overline{y}_N = 0.54420 \qquad \text{και} \qquad \sigma_N = 1.14360$$

και επομένως ο παράγοντας συχνότητας k_T είναι:

$$k_T = \frac{-\left[\ln\left(\ln\frac{T}{T-1}\right) + \overline{y}_N\right]}{\sigma_N} \Rightarrow$$

$$\Rightarrow k_T = \frac{-\left[\ln\left(\ln\frac{65}{65-1}\right) + 0.54420\right]}{1.14360} \Rightarrow k_T = 3.168$$

Μέθοδος παράγοντα συχνότητα, υπολογίζει πιο εύκολα «το αντίστροφο» πρόβλημα

Το k_T βρίσκεται αμέσως αν χρησιμοποιηθεί ο Πίνακας 2.3 ενώ ο συντελεστής διακύμανσης c_v είναι:

$$c_v = \frac{\widehat{\sigma}}{\overline{Q}} = \frac{36}{61} = 0.590$$

Τελικά από τη σχέση του παράγοντα συχνότητας έχουμε:

$$Q_T = \overline{Q}(1 + c_v \cdot k_T) = 61\,(1 + 0.590 \cdot 3.168) = 175.0 \text{ m}^3/\text{s}.$$