ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

# Exercice n° HA 0805 - Corrigé

## Estimation des débits de crue pour différents temps de retour par la méthode statistique – Application au bassin versant de la Mentue à Yvonand (VD, Suisse)

### Données de l'exercice :

L'exercice porte sur le bassin versant de la Mentue (station à Yvonand). Les données nécessaires à la réalisation de cet exercice se trouvent dans le tableau 1-énoncé ou dans le fichier Excel « exercice HA0805_enonce.xls ». Un fichier Excel « HA0805_feuilledecalcul.xls » à compléter est également disponible. Le corrigé se trouve aussi dans le fichier « HA0805_corrige.xls ».

### Question 1 : Estimation des débits de pointe de temps de retour, 5, 20, 50, 100 ans par *la méthode statistique – Ajustement de Gumbel*

⦿ Méthode à appliquer : ajustement statistique d'une série de données -Gumbel

L'objectif de cet exercice est d'estimer les débits de pointes (débits maximaux) correspondants à un certain temps de retour, c'est-à-dire à une certaine probabilité d'apparition donnée.

L'analyse fréquentielle d'une longue série de débits maximaux permet d'estimer le temps de retour d'une valeur particulière. Cette prédiction repose sur la définition et la mise en œuvre d'un modèle fréquentiel qui est une équation décrivant (modélisant) le comportement statistique d'un processus. Ces modèles décrivent la probabilité d'apparition d'un événement de valeur donnée. C'est du choix du modèle fréquentiel (et plus particulièrement de son type) que dépendra la validité des résultats de l'analyse fréquentielle.

Un modèle fréquentiel très souvent utilisé pour décrire le comportement statistique des valeurs extrêmes est la distribution statistique de Gumbel (loi double exponentielle ou loi de Gumbel). La fonction de répartition de la loi de Gumbel F(x) s'exprime de la manière suivante :

$$F(x) = \exp\left(-\exp\left(-\frac{x-a}{b}\right)\right) \quad (1) \quad \text{avec la variable réduite suivante : } u = \frac{x-a}{b} \quad (2)$$

où $a$ et $b$ sont les paramètres du modèle de Gumbel.

La distribution s'écrit alors de la manière suivante :

$$F(x) = \exp(-\exp(-u)) \quad (3) \qquad \text{et } u = -\ln(-\ln(F(x))). \quad (4)$$

L'avantage d'utiliser la variable réduite est que l'expression d'un quantile est alors linéaire ($x_q = a + bu_q$).

En conséquence, dès lors que les points de la série à ajuster peuvent être reportés dans un système d'axes $x$ - $u$ il est possible d'ajuster une droite qui passe le mieux par ces points et d'en déduire les deux paramètres $a$ et $b$ de la loi. Il existe différentes méthodes d'ajustement : méthode graphique (ajustement à l'œil ou à l'aide d'une régression statistique), méthode des moments ect.

En pratique, il s'agit essentiellement d'estimer la probabilité de non dépassement *F(xi)* qu'il convient d'attribuer à chaque valeur xi. Il existe de nombreuses formules d'estimation de la fonction de répartition $\hat{F}(x)$ à l'aide de la fréquence empirique. Elles reposent toutes sur un tri de la série par valeurs croissantes permettant d'associer à chaque valeur son rang r. Des simulations ont montré que pour la loi de Gumbel, il faut utiliser la fréquence empirique de Hazen :

$$F\left(x_{[r]}\right) = \frac{r-0.5}{n} \qquad (5)$$

où $r$ est le rang dans la série de données classée par valeurs croissantes, $n$ est la taille de l'échantillon, $x_{[r]}$ la valeur de rang $r$.

Rappelons encore que le temps de retour T d'un événement est défini comme étant l'inverse de la fréquence d'apparition de l'événement. Soit :

$$T = \frac{1}{1-F_Q(x_Q)}. \qquad (6)$$

☉ Démarche et résultats

**Etape 1** : Préparation de la série de données des débits de pointe.

- Trier les valeurs dans l'ordre croissant.
- Attribuer un rang à chaque valeur.

**Etape 2** : Calcul de la fréquence empirique pour chaque rang (Hazen, équation (5)).

**Etape 3** : Calcul de la variable réduite « u » du Gumbel (équation (4)).

**Etape 4** : Représentation graphique des couples ($u_i$, $x_i$) de la série à ajuster (figure 1).

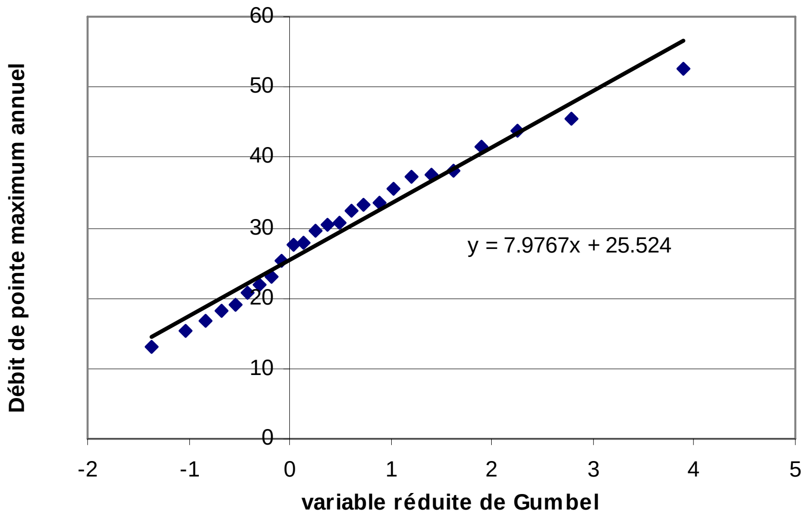


*Figure 1. Ajustement graphique du modèle*
*(calcul des paramètres « a » et « b » de la droite d'ajustement de Gumbel)*

**Etape 5** : Ajustement d'une relation linaire de type $x_q = a + bu_q$ aux couples ($u_i$, $x_i$) (figure 1) et en déduire les deux paramètres $a$ et $b$). Avec un ajustement de type graphique (à l'œil), on a alors une estimation des paramètres $a$ et $b$ :

$$a = 25.5 \text{et} \;\; b = 7.98$$

**Etape 6** : Utilisation du modèle statistique pour estimer des débits de pointe de différents temps de retour $T$. Par exemple pour $T$=100 ans, on suit les étapes suivantes :

- Calcul de la fréquence de non-dépassement d'après la relation (6) :

$$F(Q_p(T)) = 1 - \frac{1}{T} = 1 - \frac{1}{100} = 0,99$$

- Calcul de la variable réduite de Gumbel correspondante d'après la relation (4) :

$$u = -\ln\left(-\ln\left(F(Q_p(T))\right)\right) = -\ln\left(-\ln\left(0{,}99\right)\right) = 4{,}60$$

- Calcul du quantile correspondant d'après la relation linéaire (avec *a* et *b* fournis par l'étape 5 précédente ) :

$$Q_q(100) = a + bu_{100} = 25{,}57 + 7{,}98 \cdot 4{,}60 = 62{,}2 \quad \text{m}^3\text{/s}$$

On a de même pour les autres temps de retour :

$$Q_q(2{,}33) = a + bu_{2,33} = 25{,}5 + 7{,}98 \cdot 0{,}58 = 30{,}1 \quad \text{m}^3\text{/s}$$

$$Q_q(5) = a + bu_5 = 25{,}5 + 7{,}98 \cdot 1{,}50 = 37{,}5 \quad \text{m}^3\text{/s}$$

$$Q_q(20) = a + bu_{20} = 25{,}5 + 7{,}98 \cdot 2{,}97 = 49{,}2 \quad \text{m}^3\text{/s}$$

$$Q_q(50) = a + bu_{50} = 25{,}5 + 7{,}98 \cdot 3{,}90 = 56{,}6 \quad \text{m}^3\text{/s}$$

Les résultats sont regroupés dans le tableau ci-dessous :

| | **Mentue (Yvonnand)** | **Rang** | **Fréquence empirique** | **Variable réduite** | **Temps de retour** |
|---|---|---|---|---|---|
| Année | Débit [m$^3$/s]. | *r* | F(x) | u(x) | T(x) |
| 1971 | 23.00 | 8 | 0.30 | -0.186 | 1.429 |
| 1972 | 15.40 | 2 | 0.06 | -1.034 | 1.064 |
| 1973 | 13.20 | 1 | 0.02 | -1.364 | 1.020 |
| 1974 | 19.08 | 5 | 0.18 | -0.539 | 1.220 |
| 1975 | 18.09 | 4 | 0.14 | -0.676 | 1.163 |
| 1976 | 20.81 | 6 | 0.22 | -0.415 | 1.282 |
| 1977 | 41.50 | 22 | 0.86 | 1.892 | 7.143 |
| 1978 | 30.82 | 14 | 0.54 | 0.484 | 2.174 |
| 1979 | 43.67 | 23 | 0.90 | 2.250 | 10.000 |
| 1980 | 33.25 | 16 | 0.62 | 0.738 | 2.632 |
| 1981 | 27.59 | 10 | 0.38 | 0.033 | 1.613 |
| 1982 | 52.66 | 25 | 0.98 | 3.902 | 50.000 |
| 1983 | 32.55 | 15 | 0.58 | 0.607 | 2.381 |
| 1984 | 30.47 | 13 | 0.50 | 0.367 | 2.000 |
| 1985 | 37.43 | 20 | 0.78 | 1.392 | 4.545 |
| 1986 | 35.47 | 18 | 0.70 | 1.031 | 3.333 |
| 1987 | 29.65 | 12 | 0.46 | 0.253 | 1.852 |
| 1988 | 33.60 | 17 | 0.66 | 0.878 | 2.941 |
| 1989 | 16.83 | 3 | 0.10 | -0.834 | 1.111 |
| 1990 | 27.99 | 11 | 0.42 | 0.142 | 1.724 |
| 1991 | 37.27 | 19 | 0.74 | 1.200 | 3.846 |
| 1992 | 37.99 | 21 | 0.82 | 1.617 | 5.556 |
| 1993 | 25.41 | 9 | 0.34 | -0.076 | 1.515 |
| 1994 | 21.83 | 7 | 0.26 | -0.298 | 1.351 |
| 1995 | 45.40 | 24 | 0.94 | 2.783 | 16.667 |

## Question 2 : Estimation des débits de pointe de temps de retour, 5, 20, 50, 100 ans par *la méthode statistique* - Méthode des moments.

⦿ Méthode à appliquer : Méthode des moments

La méthode des moments consiste à égaler les moments des échantillons avec les moments théoriques de la loi[1]. Par la méthode des moments les paramètres $a$ et $b$ sont calculés d'après les formules :

$$\begin{cases} \hat{b} = \dfrac{\sqrt{6}}{\pi}\hat{\sigma} \\ \hat{a} = \hat{\mu} - \hat{b}\gamma. \end{cases} \quad \text{avec } \gamma = 0.5772 \text{ (constante d'Euler).}$$

avec

$\sigma$ : écart-type des valeurs composant l'échantillon.
$\mu$ : moyenne de l'échantillon.

Dès lors il est possible d'estimer les débits dont la représentation graphique est une droite d'équation :

$$\hat{Q} = \hat{a} + \hat{b} \cdot u$$

avec : $u$: variable réduite (cf. équation (4)).

⦿ Démarche et résultats

| **Paramètres** | **Débit** | **Unités** |
|---|---|---|
| Moyenne | 30 | [m$^3$/s] |
| Ecart-type | 10.13 | [m$^3$/s] |
| Paramètre *a* | 25.87 | [m$^3$/s] |
| Paramètre *b* | 7.21 | [m$^3$/s] |
| $Q_{2.33}$ | 30 | [m$^3$/s] |
| $Q_5$ | 36.7 | [m$^3$/s] |
| $Q_{20}$ | 47.3 | [m$^3$/s] |
| $Q_{50}$ | 54 | [m$^3$/s] |
| $Q_{100}$ | 59 | [m$^3$/s] |

## Question 3 : Commentaires.

Dans l'exemple ci-dessus, les deux méthodes donnent des résultats très proches l'une de l'autre. La méthode des moments est nettement plus rapide à appliquer, elle présente cependant un désavantage par rapport à la méthode graphique. L'ajustement graphique permet en effet de repérer d'éventuels points qui ne sont pas bien alignés et de ne pas en tenir compte. On pourrait également voir si la série comportait une « rupture » c'est-à-dire un changement de pente et donc un changement des paramètres de la loi statistique. De manière générale, l'ajustement manuel donne souvent beaucoup d'informations sur la série étudiée.

[1] Autrement dit : la méthode des moments consiste à égaler les caractéristiques de la distribution (empirique) des échantillons avec les caractéristiques théoriques de la loi. Les caractéristiques utilisées pour décrire une distribution sont les moments dont les plus connus sont la moyenne et la variance (la moyenne d'une distribution est appelée premier moment, la variance est le deuxième moment). Les moments d'ordre plus élevés sont utiles pour caractériser d'autres aspects de la distribution. Le troisième moment est par exemple lié à l'asymétrie ou la dyssimétrie.