

ECN 6578A, Économétrie des marchés financiers,
Hiver 2023

Cours 11 : Valeur à risque

William McCausland

2023-04-03

Pourquoi la valeur à risque, partie I

- ▶ L'objectif de l'investisseur : choisir un portefeuille selon ses préférences.
- ▶ L'information pertinente est la loi prédictive conjointe de tous les rendements des actifs.
- ▶ Gérer toute cette information est trop difficile :
 - ▶ estimation
 - ▶ élicitation des préférences
 - ▶ optimisation
- ▶ Simplification 1 : focaliser sur la moyenne et la variance des rendements : (CAPM)
- ▶ Simplification 2 : focaliser sur les grandes pertes et leur probabilité (VaR, ES)
- ▶ Pour un preneur de décisions, les deux approches sont complémentaires, pas exclusives.
- ▶ Le problème de gestion de risque pour une entreprise a des éléments en commun.

Pourquoi la valeur à risque, partie II

- ▶ Perte non-linéaire : risque de faire faillite, de devoir vendre des actifs productifs, d'avoir besoin d'un sauvetage financier
- ▶ Exemples de risque :
 - ▶ Institutions financières, levier de financement
 - ▶ Importeurs, exporteurs (risque des devises)
 - ▶ Firmes qui exploitent des ressources naturelles (risque de changements de prix)
 - ▶ Fonds de pensions : risque de perte de valeur des actifs

Valeur à Risque (VaR)

- ▶ Pour les quantités suivantes
 1. Terme (ou horizon) I (en périodes),
 2. Probabilité p de grande perte, (souvent $p = 0.01$)
 3. Fonction de répartition $F_I(\cdot)$ pour le gain de valeur d'un portefeuille en I périodes,
- ▶ la valeur à risque (VaR) est (par définition) la solution de l'équation

$$p = F_I(-\text{VaR}).$$

- ▶ la question de conditionnement
 - ▶ loi inconditionnelle, long terme, non-paramétrique
 - ▶ loi conditionnelle, court terme, besoin d'un modèle (régression quantile ou plein modèle)

RiskMetrics

- ▶ Modèle simple, IGARCH Gaussien :

$$\mu_t = 0, \quad \sigma_t^2 = \alpha\sigma_{t-1}^2 + (1 - \alpha)r_{t-1}^2, \quad r_t = a_t = \sigma_t\epsilon_t$$

où $\epsilon_t \sim N(0, 1)$.

- ▶ Un seul paramètre $\alpha \in (0, 1)$, approximativement 0.94.
- ▶ Utile à court terme, pour les rendements journaliers.
- ▶ Pour $j > i > 0$,

$$\text{cov}[r_{t+i}, r_{t+j}|F_t] = E[\sigma_{t+i}\sigma_{t+j}\epsilon_{t+i}E[\epsilon_{t+j}|F_{t+j-1}]|F_t] = 0.$$

RiskMetrics : variance multipériode

- ▶ De la diapo précédente : $\text{cov}[r_{t+i}, r_{t+j}|F_t] = 0$.
- ▶ Prochaine étape :

$$\begin{aligned}\text{Var}[a_{t+i}|F_t] &= \text{Var}[E[a_{t+i}|F_{t+i-1}]|F_t] + E[\text{Var}[a_{t+i}|F_{t+i-1}]|F_t] \\ &= E[\sigma_{t+i}^2|F_t].\end{aligned}$$

- ▶ Maintenant la variance conditionnelle (à t) du rendement k -période est

$$\sigma_t^2[k] = \sum_{i=1}^k \text{Var}[a_{t+i}|F_t] = \sum_{i=1}^k E[\sigma_{t+i}^2|F_t].$$

Une récursion pour le modèle RiskMetrics

- ▶ Pour tous t ,

$$\sigma_t^2 = \alpha\sigma_{t-1}^2 + (1 - \alpha)r_{t-i}^2 = \alpha\sigma_{t-1}^2 + (1 - \alpha)\sigma_{t-1}^2\epsilon_t^2.$$

$$\sigma_t^2 = \sigma_{t-1}^2 + (1 - \alpha)\sigma_{t-1}^2(\epsilon_{t-1}^2 - 1)$$

- ▶ En particulier,

$$\sigma_{t+i}^2 = \sigma_{t+i-1}^2 + (1 - \alpha)\sigma_{t+i-1}^2(\epsilon_{t+i-1}^2 - 1)$$

- ▶ La loi des espérances itérées donne ($E[E[\cdot|F_{t+i-2}]|F_t]$)

$$E[\sigma_{t+i}^2|F_t] = E[\sigma_{t+i-1}^2|F_t].$$

- ▶ Par induction, $E[\sigma_{t+i}^2|F_t] = \sigma_{t+1}^2$ pour chaque i , alors

$$\sigma_t^2[k] = k\sigma_{t+1}^2.$$

Calcul de VaR RiskMetrics

- ▶ À t , on détient une quantité Q d'un portefeuille à prix P_t .
- ▶ Pour $p = 0.05$, $l = 1$, la VaR est de

$$P_t Q \times -\Phi^{-1}(0.05)\sigma_{t+1} \approx P_t Q \times 1.65\sigma_{t+1}$$

- ▶ Pour $p = 0.05$, $l = k$, la VaR est de

$$P_t Q \times -\Phi^{-1}(0.05)\sqrt{k}\sigma_{t+1} \approx P_t Q \times 1.65\sqrt{k}\sigma_{t+1}$$

- ▶ Notez l'approximation $R_{t+l} = r_{t+l}$. On utilise r_{t+l} dans le modèle IGARCH, R_{t+l} pour calculer la perte.
- ▶ Exemple 7.1 :
 - ▶ $\sigma_{t+1} = 0.53\%$ (écart-type empirique du rendement journalier pour le taux d'échange DM/Dollar, juin 1997)
 - ▶ $P_t Q = 10^7$ \$
 - ▶ $p = 0.05$ alors $\Phi^{-1}(p) = \Phi^{-1}(0.05) \approx -1.65$
 - ▶ $\text{VaR}(1) = 10^7 \times 1.65 \times 0.0053 = 87450$ \$
 - ▶ $\text{VaR}(10) = 10^7 \times \sqrt{10} \times 1.65 \times 0.0053 = 276541$ \$

Discussion

- ▶ Simple
- ▶ L'hypothèse de gaussianité peut être très trompeur pour $p \leq 0.01$.

Approche économétrique I

- ▶ Un modèle ARMA(p,q)-GARCH(u,v) pour r_1, \dots, r_n :

$$\mu_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} - \sum_{j=1}^q \theta_j a_{t-j}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^u \alpha_i a_{t-i}^2 + \sum_{j=1}^v \beta_j \sigma_{t-j}^2.$$

- ▶ En évaluant la vraisemblance, on obtient les μ_t, σ_t^2 .
- ▶ Chaque μ_t et σ_t^2 est une fonction de tous les r_1, \dots, r_{t-1}
- ▶ Par la suite, on peut calculer $a_t = r_t - \mu_t, t = 1, \dots, n$.
- ▶ Sachant $(r_1, \dots, r_n), r_{n+1} \sim (\mu_{n+1}, \sigma_{n+1})$, où

$$\mu_{n+1} = \phi_0 + \sum_{i=1}^p \phi_i r_{n+1-i} - \sum_{j=1}^q \theta_j a_{n+1-j}$$

$$\sigma_{n+1}^2 = \alpha_0 + \sum_{i=1}^u \alpha_i a_{n+1-i}^2 + \sum_{j=1}^v \beta_j \sigma_{n+1-j}^2.$$

Approche économétrique II

- ▶ Si $r_{n+1}|r_1, \dots, r_n \sim N(\mu_{n+1}, \sigma_{n+1}^2)$, la VaR pour une période avec $p = 0.05$ est, par unité de valeur et avec l'approximation $R_{t+1} = r_{t+1}$,

$$-\text{VaR} = \mu_{n+1} + \Phi^{-1}(0.05)\sigma_{n+1} = \mu_{n+1} - 1.65\sigma_{n+1}.$$

- ▶ Si $r_{n+1}|r_1, \dots, r_n \sim t(\mu_{n+1}, \sigma_{n+1}^2, \nu)$:
 - ▶ La variance d'une variable aléatoire t de Student standard avec ν degrés de liberté est $\nu/(\nu - 2)$.
 - ▶ Alors si $\text{Var}[\lambda Z] = \sigma^2$ et $Z \sim t_\nu$, $\lambda = \sigma/\sqrt{\nu/(\nu - 2)}$.
 - ▶ Soit $t_\nu(p)$ la quantile d'une variable aléatoire $t(\nu)$.
 - ▶ La valeur à risque est, par unité de valeur

$$-\text{VaR} = \mu_{n+1} + \frac{t_\nu(p)\sigma_{n+1}}{\sqrt{\nu/(\nu - 2)}}.$$

Exemple, calcul du VaR ou $r_{n+1} | r_1, \dots, r_n$ est un t de Student

- ▶ Mettons que $\mu_{n+1} = 0.001$, $\sigma_{n+1} = 0.02$, $\nu = 12$.
- ▶ Exemple, calcul du VaR, par unité de valeur, pour $p = 0.05$

```
p = 0.05
nu = 12
mu.np1 = 0.001
sigma.np1 = 0.02
t.nu.p = qt(p, nu)
mVaR = mu.np1 + t.nu.p * sigma.np1 / (sqrt(nu/(nu-2)))
-mVaR

## [1] 0.03153997
```

Estimation quantile, approche inconditionnelle

- ▶ Trier les rendements r_1, \dots, r_n pour calculer les statistiques d'ordre :

$$r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(n)}.$$

- ▶ Supposons que r_t est iid avec fonction de répartition F et densité f .
- ▶ On veut estimer $x_p = F^{-1}(p)$, la quantile p de la population
- ▶ Pour $l = np$ entier, $r_{(l)}$ est la quantile p de l'échantillon et

$$r_{(l)} \sim_{\text{asy}} N \left(x_p, \frac{p(1-p)}{n[f(x_p)]^2} \right).$$

- ▶ $r_{(l)}$ est une estimation de x_p .
- ▶ Il faut estimer $f(x_p)$ pour calculer la variance de l'estimateur.

Quand np n'est pas entier

- ▶ Trouver l_1, l_2 entiers tels que $l_2 = l_1 + 1, l_1 < np < l_2$.
- ▶ Alors l_1 est le plancher de np : $\lfloor np \rfloor$.
- ▶ Soit $p_1 = l_1/n, p_2 = l_2/n$.
- ▶ Trouver \hat{x}_p entre \hat{x}_{p_1} et \hat{x}_{p_2} :

$$\hat{x}_p = \frac{p_2 - p}{p_2 - p_1} r_{(l_1)} + \frac{p - p_1}{p_2 - p_1} r_{(l_2)}$$

Exemple à long terme

```
# Séries IBM journalière
r = scan('d-ibmln.txt')
r.sort = sort(r)
n = length(r)
# Calcul de x-chapeau-p
p = 0.05
l1 = floor(p*n); p1 = l1/n;
l2 = ceiling(p*n); p2 = l2/n;
x.ch.p = ((p2-p) * r.sort[l1] + (p-p1) * r.sort[l2]) / (p2-
# Calcul de f(x-chapeau-p)
ds = density(r)
x.ch.p.index = findInterval(x.ch.p, ds$x)
f.x.ch.p = ds$y[x.ch.p.index]
sigma = sqrt(p*(1-p)/n)/f.x.ch.p
```

Valeurs

```
p1;p2;l1;l2
```

```
## [1] 0.04989931
```

```
## [1] 0.05001119
```

```
## [1] 446
```

```
## [1] 447
```

```
x.ch.p;f.x.ch.p;sigma
```

```
## [1] -2.1492
```

```
## [1] 0.05977387
```

```
## [1] 0.03856694
```


Illustration : histogramme

```
hist(r, 80)  
abline(v=x.ch.p)
```

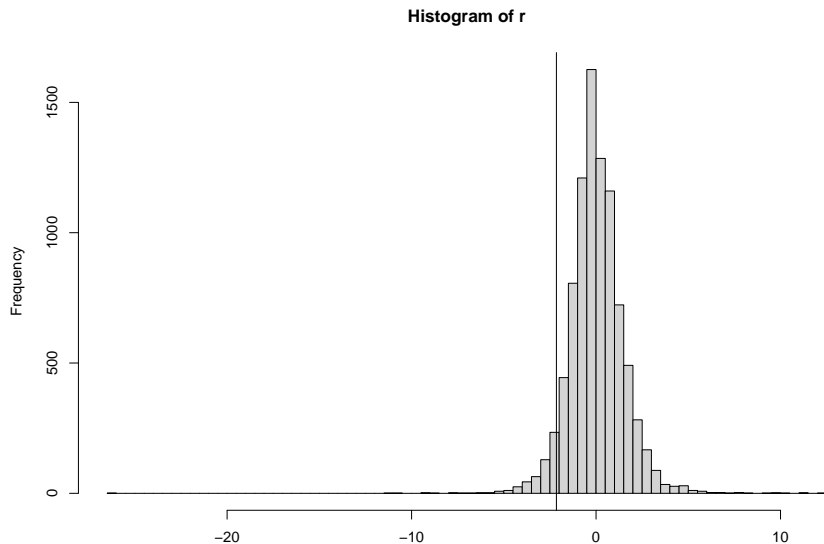
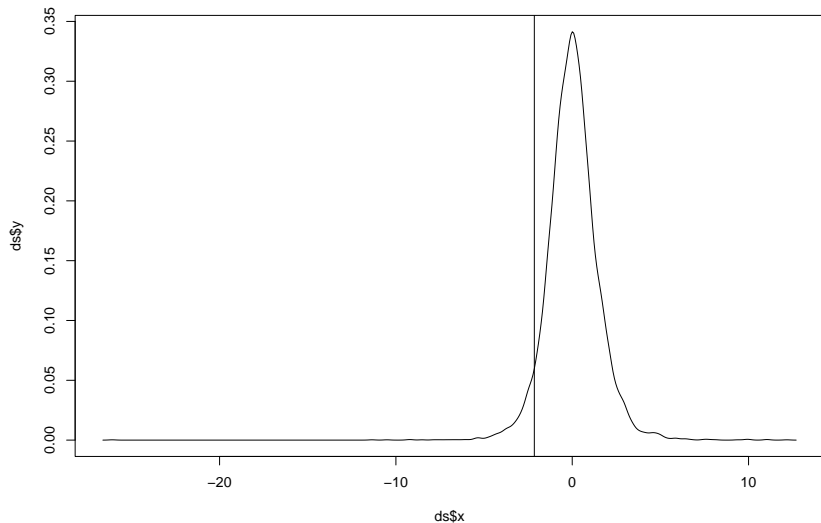


Illustration : densité

```
plot(ds$x, ds$y, type='l')  
abline(v=x.ch.p)
```



Commentaires

▶ Avantages :

1. Simple
2. Pas de modèle
3. Plus de précision dans les ailles

▶ Inconvénients :

1. L'hypothèse de iid peu crédible : quand la volatilité est persistante, la variance asymptotique de la quantile de l'échantillon est plus grande.
2. Pas de conditionnement à l'information pertinente.
3. Les quantiles empiriques ne sont pas précises pour p petit.
4. Il faut supposer qu'il n'y aie pas de changement de loi entre la période de l'échantillon et la période de prévision.

Aparté : la moyenne comme choix optimal

- ▶ Soit r une variable aléatoire d'une loi, r_1, \dots, r_n un échantillon de cette loi.
- ▶ On veut choisir une valeur représentative de la loi ou de l'échantillon.
- ▶ La moyenne $\beta = E[r]$ est la solution du problème suivant de minimization de perte quadratique espérée :

$$\arg \min_{\beta} E[(r - \beta)^2] = \arg \min_{\beta} E[(r - E[r])^2] + (E[r] - \beta)^2.$$

- ▶ La moyenne de l'échantillon $b = \bar{r} = n^{-1} \sum_{i=1}^n r_i$ est la solution du problème de moindres carrés:

$$\arg \min_b n^{-1} \sum_{i=1}^n (r_i - b)^2.$$

Aparté : les quantiles comme choix optimaux

- ▶ La quantile $x_p = F^{-1}(p)$ est la solution du problème suivant minimization de perte espérée :

$$x_p = \arg \min_{\beta} E[w_p(r - \beta)],$$

- ▶ La fonction de perte w_p est

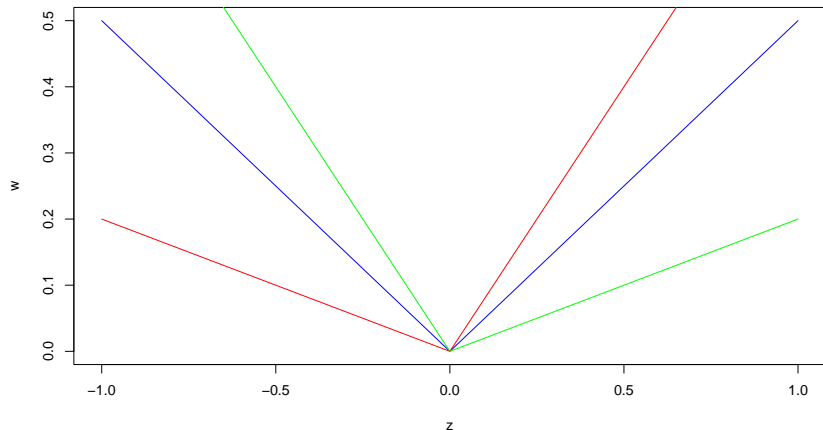
$$w_p(z) = \begin{cases} pz, & z \geq 0 \\ -(1-p)z, & z < 0. \end{cases}$$

- ▶ La quantile empirique (ou de l'échantillon) est la solution de

$$\hat{x}_p = \arg \min_{\beta} \sum_{i=1}^n w_p(r_i - \beta).$$

La fonction $w_p(z)$

```
z = seq(-1, 1, by=0.1)
p=0.5; plot(z, p*z*(z>=0) - (1-p)*z*(z<0), col='blue', type='l')
p=0.8; lines(z, p*z*(z>=0) - (1-p)*z*(z<0), col='red')
p=0.2; lines(z, p*z*(z>=0) - (1-p)*z*(z<0), col='green')
```



Régression quantile

- ▶ Supposons que la quantile conditionnelle de r_t sachant x_t , dénoté $q_p|x_t$, est linéaire en x_t : $q_p|x = \beta_p^\top x$.
- ▶ Alors β_p est la solution de

$$\beta_p = \arg \min_{\beta} E[w_p(r - \beta^\top x)].$$

- ▶ L'analogie dans l'échantillon donne l'estimateur

$$\hat{\beta}_p = \arg \min_{\beta} \sum_{t=1}^n w_p(r_t - \beta^\top x_t).$$

- ▶ Notes

- ▶ Dans le contexte de VaR, x_t comprend des variables dans F_{t-1} .
- ▶ la fonction de perte w_p est moins sensible aux valeurs aberrantes que la fonction de perte quadratique.