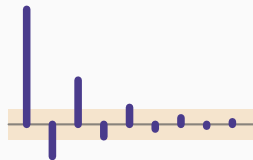


Modélisation ARMA des rentabilités financières

Chapitre 16 : La prévision — ponctuelle, par intervalle, par densité



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Avant de commencer

La prévision ponctuelle

La prévision par intervalle

La prévision par densité

La limite qui appelle le cours suivant

Synthèse

Avant de commencer

Trois réponses, de plus en plus riches

Ponctuelle : un nombre. « La rentabilité de demain sera de 0,04 %. »

Par intervalle : une fourchette et sa probabilité. « Entre $-2,3\%$ et $+2,4\%$ avec 95 % de chances. »

Par densité : toute la loi de y_{T+h} . C'est la réponse complète — les deux autres n'en sont que des résumés.

En finance, c'est la troisième qui compte

Une VaR est un **quantile** de la densité prévue. Le prix d'une option est une **espérance** sous cette densité. La prévision ponctuelle, elle, n'intéresse presque personne sur données de marché.

La prévision ponctuelle

Meilleure prévision au sens de l'erreur quadratique

La prévision $\hat{y}_{T+h}$ qui minimise $E[(y_{T+h} - \hat{y}_{T+h})^2 \mid \Omega_T]$ est l'**espérance conditionnelle**

$$\hat{y}_{T+h|T} = E(y_{T+h} \mid \Omega_T).$$

Un résultat, pas une convention

Ce n'est pas un choix parmi d'autres : sous le critère quadratique, aucune fonction du passé ne fait mieux. Changer de critère changerait la réponse — la médiane conditionnelle minimiserait l'erreur absolue.

Trois substitutions, appliquées à l'équation du modèle

Pour toute date future $T + h$, on remplace :

- y_{T+j} **connu** ($j \leq 0$) par sa valeur observée;
- y_{T+j} **futur** ($j > 0$) par sa propre prévision $\hat{y}_{T+j|T}$;
- ε_{T+j} **futur** ($j > 0$) par **zéro**; ε_{T+j} passé par le résidu $\hat{\varepsilon}_{T+j}$.

Pourquoi zéro

Un choc futur est par définition imprévisible et centré : sa meilleure prévision est son espérance, qui vaut zéro. C'est là que se loge toute l'incertitude de la prévision.

Avec $y_t = \theta + \alpha y_{t-1} + \varepsilon_t$:

$$\hat{y}_{T+1|T} = \theta + \alpha y_T$$

$$\hat{y}_{T+2|T} = \theta + \alpha \hat{y}_{T+1|T} = \theta(1 + \alpha) + \alpha^2 y_T$$

$$\hat{y}_{T+h|T} = \mu + \alpha^h (y_T - \mu), \quad \mu = \frac{\theta}{1 - \alpha}.$$

La convergence vers la moyenne

Comme $|\alpha| < 1$, $\hat{y}_{T+h|T} \rightarrow \mu$ quand $h \rightarrow \infty$. **Au-delà de quelques périodes, un modèle stationnaire ne prévoit plus rien d'autre que la moyenne.**

Un ordre de grandeur

$\alpha = 0,6$, $y_T - \mu = 10$. L'écart prévu vaut 6 à un pas, 3,6 à deux pas, 0,6 à cinq pas, 0,06 à dix pas. La mémoire du modèle s'épuise en une poignée de périodes.

$$\hat{y}_{T+1|T} = \theta + \beta_1 \hat{\varepsilon}_T + \dots + \beta_q \hat{\varepsilon}_{T+1-q}$$

Un horizon strictement borné

Pour $h > q$, tous les chocs concernés sont futurs, donc remplacés par zéro :

$$\hat{y}_{T+h|T} = \theta = \mu \quad (h > q).$$

La mémoire finie se voit dans la prévision

Un MA(2) prévoit quelque chose à un et deux pas, et **strictement rien** au-delà. La coupure de l'ACF au retard q et la coupure de la prévision à l'horizon q sont le même phénomène.

La prévision par intervalle

Décomposition

En représentation MA(∞), $y_t = \mu + \sum_{j \geq 0} \psi_j \varepsilon_{t-j}$, l'erreur à l'horizon h s'écrit

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T} = \sum_{j=0}^{h-1} \psi_j \varepsilon_{T+h-j}.$$

Elle ne contient **que** des chocs futurs — ceux qu'on ne pouvait pas connaître.

$$V(e_{T+h}) = \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \psi_j^2.$$

Trois enseignements

Elle **croît** avec h : chaque période supplémentaire ajoute un choc.

Elle **converge** vers $\sigma_\varepsilon^2 \sum_{j \geq 0} \psi_j^2 = V(y_t)$, la variance inconditionnelle. À long horizon, prévoir ne réduit plus l'incertitude du tout.

À un pas, elle vaut σ_ε^2 : c'est le **minimum absolu**, l'incertitude irréductible.

AR(1) : la formule explicite

$\psi_j = \alpha^j$, donc

$$V(e_{T+h}) = \sigma_\varepsilon^2 \frac{1 - \alpha^{2h}}{1 - \alpha^2} \xrightarrow{h \rightarrow \infty} \frac{\sigma_\varepsilon^2}{1 - \alpha^2} = \gamma_0.$$

Intervalle de prévision à 95 pour cent

$$\hat{y}_{T+h|T} \pm 1,96 \sqrt{V(e_{T+h})}.$$

Trois hypothèses, souvent oubliées

Cet intervalle suppose que les innovations sont **gaussiennes**, que les paramètres sont **connus** (et non estimés), et que le modèle est **bien spécifié**. Chacune est fausse en pratique; l'intervalle réel est plus large que l'intervalle affiché.

Intervalle de prévision, intervalle de confiance

Ne pas confondre. Un intervalle de **confiance** encadre un paramètre inconnu mais fixe. Un intervalle de **prévision** encadre une variable aléatoire non encore réalisée : il est nécessairement plus large, puisqu'il intègre l'aléa futur en plus de l'incertitude d'estimation.

La prévision par densité

Densité prédictive

$$f(y_{T+h} \mid \Omega_T).$$

Sous hypothèse gaussienne et paramètres connus :

$$y_{T+h} \mid \Omega_T \sim \mathcal{N}\left(\hat{y}_{T+h|T}, \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \psi_j^2\right).$$

Ce que la densité apporte de plus

Elle donne **n'importe quel** quantile, donc la VaR; la probabilité d'un événement, par exemple $\mathbb{P}(y_{T+h} < 0)$; et l'espérance de toute fonction du futur, ce dont l'évaluation d'options a besoin.

Densité par simulation

1. Tirer N trajectoires de chocs futurs $(\varepsilon_{T+1}, \dots, \varepsilon_{T+h})$.
2. Propager le modèle sur chacune pour obtenir $y_{T+h}^{(i)}$.
3. La distribution empirique des $y_{T+h}^{(i)}$ **est** la densité prédictive.

Les deux façons de tirer les chocs

Paramétrique : tirer dans $\mathcal{N}(0, \hat{\sigma}_\varepsilon^2)$ — simple, mais impose la normalité.

Bootstrap : rééchantillonner dans les **résidus observés** — conserve l'asymétrie et les queues épaisses, sans hypothèse de loi. C'est la voie à privilégier en finance.

La transformée en probabilité intégrale

Si la densité prévue est correcte, les valeurs $u_t = F_t(y_t)$ — la fonction de répartition prévue évaluée en la réalisation — sont **uniformes sur $[0, 1]$ et indépendantes**.

Le lien avec la VaR

C'est exactement le principe du *backtesting* réglementaire : compter les dépassements de VaR revient à vérifier que u_t franchit le seuil 1 % dans 1 % des cas, et sans se regrouper dans le temps. Le cours de Value at Risk y consacre son chapitre 6.

La limite qui appelle le cours
suivant

Ce qui est juste et ce qui ne l'est pas

Le diagnostic, sur données financières

La prévision **ponctuelle** est correcte — et sans intérêt : elle vaut à peu près la moyenne dès le premier pas, puisqu'il n'y a presque pas de mémoire.

La prévision par **intervalle** et par **densité** est fausse, parce qu'elle suppose une variance constante que le chapitre 15 a montrée changeante.

Le mécanisme de l'erreur

En période calme, l'intervalle fondé sur $\hat{\sigma}_\varepsilon^2$ — la volatilité **moyenne** de l'échantillon — est trop large : il ne sert à rien.

En période agitée, il est beaucoup trop étroit : il sous-estime le risque **précisément au moment où l'on en a besoin**.

Ce que les GARCH viennent corriger

$$\hat{y}_{T+h|T} \pm 1,96 \sqrt{\hat{h}_{T+h|T}} \quad \text{au lieu de} \quad \hat{y}_{T+h|T} \pm 1,96 \hat{\sigma}_\varepsilon.$$

La prévision ponctuelle est inchangée. C'est la **largeur** de l'intervalle qui devient conditionnelle au passé récent — et c'est tout l'objet du cours suivant.

Synthèse

L'essentiel du chapitre 16

- La meilleure prévision au sens quadratique est l'espérance conditionnelle $E(y_{T+h} \mid \Omega_T)$.
- Règle de calcul : chocs futurs remplacés par **zéro**, valeurs futures par leurs prévisions.
- AR : $\hat{y}_{T+h|T} = \mu + \alpha^h(y_T - \mu)$, convergence vers μ . MA(q) : prévision égale à μ dès $h > q$.
- $V(e_{T+h}) = \sigma_\varepsilon^2 \sum_{j < h} \psi_j^2$: croissante en h , bornée par la variance inconditionnelle.
- Densité prédictive : gaussienne sous hypothèse forte, **bootstrap des résidus** sinon.
- Sur données financières, le point est bon mais l'intervalle est faux : c'est la variance conditionnelle qu'il faut modéliser.

Vers les chapitres 17 et 18

Toute la démarche est en place. Il reste à l'exécuter — sur un vrai jeu de données, avec un vrai logiciel. R d'abord, Python ensuite.