Concentration Inequalities and Laws of Large Numbers under Epistemic Irrelevance

Fabio Gagliardi Cozman
Universidade de São Paulo
fgcozman@usp.br

1 Introduction

Goal of this paper: to derive concentration inequalities and laws of large numbers under weak assumptions of irrelevance, expressed through lower and upper expectations (closely related to De Cooman and Miranda's recent results).

2 Definitions and Assumptions

- Given a set of expectation functionals, lower and upper expectations of variable X are respectively
$$\underline{E}[X] = \inf E[X], \quad \overline{E}[X] = \sup E[X].$$
- Epistemic irrelevance of variables $X_1, \ldots, X_{i-1}$ to X_i obtains when:
$$\overline{E}[f(X_i)|A(X_{1:i-1})] = \overline{E}[f(X_i)]$$
for any bounded function f of X_i and any nonempty event $A(X_{1:i-1})$ defined by variables $X_{1:i-1}$.
(Notation: $X_{1:i}$ for $X_1, \ldots, X_i$.)
- Forward irrelevance: for each $i \in [2, n]$, variables $X_1, \ldots, X_{i-1}$ are epistemically irrelevant to X_i.
- Weak forward irrelevance: for each $i \in [2, n]$ and any nonempty event $A(X_{1:i-1})$,
$$\underline{E}[X_i|A(X_{1:i-1})] = \underline{E}[X_i],$$
$$\overline{E}[X_i|A(X_{1:i-1})] = \overline{E}[X_i].$$
- Results assume disintegrability:
$$\overline{E}[W] \leq \overline{E}\left[\overline{E}[W|Z]\right]$$
for any $W \geq 0$, $Z \geq 0$ of interest (W and Z may be sets of variables!).
Under disintegrability: if $f_i(X_i) \geq 0$ for $i \in \{1, \ldots, n\}$, then
$$\overline{E}\left[\prod_{i=1}^n f_i(X_i)\right] \leq \overline{E}\left[\ldots \overline{E}\left[\overline{E}\left[\prod_{i=1}^n f_i(X_i)|X_{1:n-1}\right]|X_{1:n-2}\right]\ldots\right];$$
so for bounded and nonnegative functions, forward irrelevance implies
$$\overline{E}\left[\prod_{i=1}^n f_i(X_i)\right] \leq \prod_{i=1}^n \overline{E}[f_i(X_i)]. \quad (1)$$

3 Bounded variables

Assume $|X_i| \leq B_i$; define
$$\gamma_n \doteq \sum_{i=1}^n B_i^2 > 0.$$

Theorem 1 (Hoeffding-like) *If bounded variables $X_1, \ldots, X_n$ satisfy Expression (1), then if $\gamma_n > 0$,*
$$\overline{P}\left(\sum_{i=1}^n (X_i - \overline{E}[X_i]) \geq \epsilon\right) \leq e^{-2\epsilon^2/\gamma_n},$$
$$\overline{P}\left(\sum_{i=1}^n (X_i - \underline{E}[X_i]) \leq -\epsilon\right) \leq e^{-2\epsilon^2/\gamma_n}.$$

Theorem 2 (Azuma-like) *If bounded variables $X_1, \ldots, X_n$ satisfy weak forward irrelevance and disintegrability holds, then if $\gamma_n > 0$,*
$$\overline{P}\left(\sum_{i=1}^n (X_i - \overline{E}[X_i]) \geq \epsilon\right) \leq e^{-2\epsilon^2/\gamma_n},$$
$$\overline{P}\left(\sum_{i=1}^n (X_i - \underline{E}[X_i]) \leq -\epsilon\right) \leq e^{-2\epsilon^2/\gamma_n}.$$

From these, get analogues to De Cooman and Miranda's results. Define
$$\underline{\mu}_n \doteq \frac{\sum_{i=1}^n \underline{E}[X_i]}{n}, \quad \overline{\mu}_n \doteq \frac{\sum_{i=1}^n \overline{E}[X_i]}{n}.$$

Theorem 3 *If bounded variables $X_1, \ldots, X_n$ satisfy weak forward irrelevance and disintegrability holds, then for any $\epsilon > 0$,*
$$\underline{P}\left(\underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) \geq 1 - 2e^{\frac{-2n\epsilon^2}{(\max_i B_i)^2}},$$
and there is N such that for any N',
$$\underline{P}\left(\forall n \in [N, N+N'] : \underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) > 1 - 2\epsilon.$$

4 Unbounded variables

Assume elementwise countable additivity (implies elementwise disintegrability).
Define:
$$Y_n \doteq \sum_{i=1}^n X_i - E_P[X_i|X_{1:i-1}].$$
Sequence $\{Y_n\}$ is a martingale; from this:
$$E_P[Y_n^2] = \sum_{i=1}^n E_P\left[(X_i - E_P[X_i|X_{1:i-1}])^2\right].$$

Using this expression and the Kolmogorov-Hajek-Renyi inequality (for martingales):

Theorem 4 *Assume elementwise countable additivity. If variables $X_1, \ldots, X_n$ satisfy weak forward irrelevance, and $\underline{E}[X_i]$ and $\overline{E}[X_i]$ are finite quantities such that $\overline{E}[X_i] - \underline{E}[X_i] \leq \delta$, and the variance of any X_i is no larger than a finite quantity σ^2, then for any $\epsilon > 0$,*
$$\underline{P}\left(\underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) \geq 1 - \frac{\sigma^2 + \delta^2}{\epsilon^2 n},$$
and there is $N > 0$ such that for any $N' > 0$,
$$\underline{P}\left(\forall n \in [N, N+N'] : \underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) > 1 - 2\epsilon.$$
Consequently,
$$\forall \epsilon > 0: \lim_{n\to\infty} \underline{P}\left(\underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) = 1,$$
$$\underline{P}\left(\limsup_{n\to\infty} \left(\frac{\sum_{i=1}^n X_i}{n} - \overline{\mu}_n\right) \leq 0\right) = 1,$$
$$\underline{P}\left(\liminf_{n\to\infty} \left(\frac{\sum_{i=1}^n X_i}{n} - \underline{\mu}_n\right) \geq 0\right) = 1.$$

5 Discussion

- For bounded variables, results are similar to De Cooman and Miranda's (slightly sharper inequalities, with proofs closer to standard results).
- Main contribution: use of martingales to handle unbounded variables.
- In fact, it seems that epistemic irrelevance is closer to martingales than to stochastic independence (to explore...).
- Also, the role of disintegrability needs further clarification. Disintegrability...
  - holds in finite spaces, or
  - holds under countable additivity, or
  - can be imposed as rationality axiom.
- Final note: There are several ways to handle conditioning on events of zero probability: results hold for full conditional measures and for regular conditioning.

The author is partially supported by CNPq. Thanks to Gert de Cooman and Enrique Miranda for their generous suggestions regarding content and presentation.

Escola Politécnica (Engineering School), Universidade de São Paulo. Address: Av. Prof. Mello Moraes, 2231 - Cidade Universitária CEP 05508900 - São Paulo, SP - Brazil