Concentration Inequalities and Laws of Large Numbers under Epistemic Irrelevance

Fabio Gagliardi Cozman

Universidade de São Paulo

fgcozman@usp.br

A bit about the speaker...

- From Brazil, took PhD at Carnegie Mellon, US.
  - Main duty during PhD: captain of the Viking Death Rats.

- Back to Brazil, now Professor at Universidade de São Paulo.
- Interested in credal sets and in models represented by graphs, where independence is important.

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.

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.

- Concentration inequalities: probability that average is far from expectation is exponentially small in the number of observations.
- Laws of large numbers: an average converges to the expectation as the number of observation increases.

Assumptions

- Forward irrelevance: for each $i \in [2, n]$, for any bounded function f of X_i and any nonempty event $A(X_{1:i-1})$,

$$\underline{E}[f(X_i)|A(X_{1:i-1})] = \underline{E}[f(X_i)]\,,$$

$$\overline{E}[f(X_i)|A(X_{1:i-1})] = \overline{E}[f(X_i)]\,.$$

- Weak forward irrelevance: for each $i \in [2, n]$ and any nonempty event $A(X_{1:i-1})$,

$$\underline{E}[X_i|A(X_{1:i-1})] = \underline{E}[X_i]\,,$$

$$\overline{E}[X_i|A(X_{1:i-1})] = \overline{E}[X_i]\,.$$

More Assumptions

- 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!).

- Finite spaces, countable additivity, rationality axiom...

Bounded variables

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

Theorem 1 (Hoeffding-like). *If bounded variables* $X_1, \ldots, X_n$ *satisfy forward irrelevance and disintegrability holds, then if* $\gamma_n > 0$,

$$\overline{P}\left(\sum_{i=1}^n (X_i - \overline{E}[X_i]) \geq \epsilon\right) \leq e^{-2\epsilon^2/\gamma_n},$$

$$\overline{P}\left(\sum_{i=1}^n (X_i - \underline{E}[X_i]) \leq -\epsilon\right) \leq e^{-2\epsilon^2/\gamma_n}.$$

Another Result for Bounded Variables

Theorem 2 (Azuma-like). *If bounded variables $X_1, \ldots, X_n$ satisfy weak forward irrelevance and disintegrability holds, then if $\gamma_n > 0$,*

$$\overline{P}\left(\sum_{i=1}^{n}(X_i - \overline{E}[X_i]) \geq \epsilon\right) \leq e^{-2\epsilon^2/\gamma_n},$$

$$\overline{P}\left(\sum_{i=1}^{n}(X_i - \underline{E}[X_i]) \leq -\epsilon\right) \leq e^{-2\epsilon^2/\gamma_n}.$$

These bounds are slightly sharper than De Cooman and Miranda's bounds. The proofs of these inequalities are interesting.

Laws of Large Numbers

From these inequalities, 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',

$$\underline{P}\left(\forall n \in [N, N+N'] : \underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) > 1 - 2\epsilon.$$

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; using this fact and the Kolmogorov-Hajek-Renyi inequality (for martingales), we get our main result.

Main Result

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* $N' > 0$*,*
$\underline{P}\left(\forall n \in [N, N + N'] : \underline{\mu}_n - \epsilon < \frac{\sum_{i=1}^n X_i}{n} < \overline{\mu}_n + \epsilon\right) > 1 - 2\epsilon$*.*

(We can take limits, given we are assuming countable additivity!)

Discussion

- Main contribution: use of martingales to handle unbounded variables.
  - In fact, it seems that epistemic irrelevance is closer to martingales than to stochastic independence (something to explore...).
- The role of disintegrability needs further clarification.
- Final note: paper deals with several ways to define conditioning.