

Negatively Associated Random Variables

Emin Karayel

February 6, 2026

Abstract

Negative Association is a generalization of independence for random variables, that retains some of the key properties of independent random variables. In particular closure properties, such as composition with monotone functions, as well as, the well-known Chernoff-Hoeffding bounds.

This entry introduces the concept and verifies the most important closure properties, as well as, the concentration inequalities. It also verifies the FKG inequality, which is a generalization of Chebyshev's sum inequality for distributive lattices and a key tool for establishing negative association, but has also many applications beyond the context of negative association, in particular, statistical physics and graph theory.

As an example, permutation distributions are shown to be negatively associated, from which many more sets of negatively random variables can be derived, such as, e.g., n -subsets, or the the balls-into-bins process.

Finally, the entry derives a correct false-positive rate for Bloom filters using the library.

Contents

1	Preliminary Definitions and Lemmas	2
2	Definition	7
3	Chernoff-Hoeffding Bounds	12
4	The FKG inequality	15
5	Preliminary Results on Lattices	17
6	Permutation Distributions	20
7	Application: Bloom Filters	26

1 Preliminary Definitions and Lemmas

```
theory Negative-Association-Util
imports
  Concentration-Inequalities.Concentration-Inequalities-Preliminary
  Universal-Hash-Families.Universal-Hash-Families-More-Product-PMF
begin
```

```
abbreviation (input) flip ::  $\langle 'a \Rightarrow 'b \Rightarrow 'c \rangle \Rightarrow 'b \Rightarrow 'a \Rightarrow 'c$  where
   $\langle flip\ f\ x\ y \equiv f\ y\ x \rangle$ 
```

Additional introduction rules for boundedness:

```
lemma bounded-const-min:
  fixes  $f :: 'a \Rightarrow real$ 
  assumes bdd-below ( $f \text{ ' } M$ )
  shows bounded ( $(\lambda x. \min\ c\ (f\ x)) \text{ ' } M$ )
   $\langle proof \rangle$ 
```

```
lemma bounded-prod:
  fixes  $f :: 'i \Rightarrow 'a \Rightarrow real$ 
  assumes finite  $I$ 
  assumes  $\bigwedge i. i \in I \implies bounded\ (f\ i\ \text{ ' } T)$ 
  shows bounded ( $(\lambda x. (\prod\ i \in I. f\ i\ x)) \text{ ' } T$ )
   $\langle proof \rangle$ 
```

```
lemma bounded-vec-mult-comp:
  fixes  $f\ g :: 'a \Rightarrow real$ 
  assumes bounded ( $f \text{ ' } T$ ) bounded ( $g \text{ ' } T$ )
  shows bounded ( $(\lambda x. (f\ x) *R\ (g\ x)) \text{ ' } T$ )
   $\langle proof \rangle$ 
```

```
lemma bounded-max:
  fixes  $f :: 'a \Rightarrow real$ 
  assumes bounded ( $(\lambda x. f\ x) \text{ ' } T$ )
  shows bounded ( $(\lambda x. \max\ c\ (f\ x)) \text{ ' } T$ )
   $\langle proof \rangle$ 
```

```
lemma bounded-of-bool: bounded (range of-bool)  $\langle proof \rangle$ 
```

```
lemma bounded-range-imp:
  assumes bounded (range  $f$ )
  shows bounded ( $(\lambda \omega. f\ (h\ \omega)) \text{ ' } S$ )
   $\langle proof \rangle$ 
```

The following allows to state integrability and conditions about the integral simultaneously, e.g. *has-int-that* $M\ f\ (\lambda x. x \leq c)$ says f is integrable on M and the integral smaller or equal to c .

```
definition has-int-that where
```

$has-int-that\ M\ f\ P = (integrable\ M\ f \wedge (P\ (\int\ \omega.\ f\ \omega\ \partial M)))$

lemma *true-eq-iff*: $P \implies True = P$ $\langle proof \rangle$

lemma *le-trans*: $y \leq z \implies x \leq y \longrightarrow x \leq (z :: 'a :: order)$ $\langle proof \rangle$

lemma *has-int-that-mono*:

assumes $\bigwedge x. P\ x \longrightarrow Q\ x$

shows $has-int-that\ M\ f\ P \leq has-int-that\ M\ f\ Q$
 $\langle proof \rangle$

lemma *has-int-thatD*:

assumes $has-int-that\ M\ f\ P$

shows $integrable\ M\ f\ P\ (integral^L\ M\ f)$
 $\langle proof \rangle$

This is useful to specify which components a functional depends on.

definition *depends-on* :: $(('a \Rightarrow 'b) \Rightarrow 'c) \Rightarrow 'a\ set \Rightarrow bool$

where $depends-on\ f\ I = (\forall x\ y. restrict\ x\ I = restrict\ y\ I \longrightarrow f\ x = f\ y)$

lemma *depends-onI*:

assumes $\bigwedge x. f\ x = f\ (\lambda i. if\ i \in I\ then\ (x\ i)\ else\ undefined)$

shows $depends-on\ f\ I$
 $\langle proof \rangle$

lemma *depends-on-comp*:

assumes $depends-on\ f\ I$

shows $depends-on\ (g \circ f)\ I$
 $\langle proof \rangle$

lemma *depends-on-comp-2*:

assumes $depends-on\ f\ I$

shows $depends-on\ (\lambda x. g\ (f\ x))\ I$
 $\langle proof \rangle$

lemma *depends-onD*:

assumes $depends-on\ f\ I$

shows $f\ \omega = f\ (\lambda i \in I. (\omega\ i))$
 $\langle proof \rangle$

lemma *depends-onD2*:

assumes $depends-on\ f\ I\ restrict\ x\ I = restrict\ y\ I$

shows $f\ x = f\ y$
 $\langle proof \rangle$

lemma *depends-on-empty*:

assumes $depends-on\ f\ \{\}$

shows $f\ \omega = f\ undefined$
 $\langle proof \rangle$

lemma *depends-on-mono*:
assumes $I \subseteq J$ *depends-on* $f I$
shows *depends-on* $f J$
 \langle *proof* \rangle

abbreviation *square-integrable* $M f \equiv$ *integrable* $M ((\text{power2} :: \text{real} \Rightarrow \text{real}) \circ f)$

There are many results in the field of negative association, where a statement is true for simultaneously monotone or anti-monotone functions. With the below construction, we introduce a mechanism where we can parameterize on the direction of a relation:

datatype *RelDirection* = *Fwd* | *Rev*

definition *dir-le* :: *RelDirection* \Rightarrow ($'d :: \text{order}$) \Rightarrow ($'d :: \text{order}$) \Rightarrow *bool* (**infixl** $\leq_{\geq 1}$ 60)
where *dir-le* $\eta =$ (if $\eta =$ *Fwd* then (\leq) else (\geq))

lemma *dir-le[simp]*:
 $(\leq_{Fwd}) = (\leq)$
 $(\leq_{Rev}) = (\geq)$
 \langle *proof* \rangle

definition *dir-sign* :: *RelDirection* \Rightarrow $'a :: \{\text{one}, \text{uminus}\}$ (± 1)
where *dir-sign* $\eta =$ (if $\eta =$ *Fwd* then 1 else (-1))

lemma *dir-le-refl*: $x \leq_{\eta} x$
 \langle *proof* \rangle

lemma *dir-sign[simp]*:
 $(\pm_{Fwd}) = (1)$
 $(\pm_{Rev}) = (-1)$
 \langle *proof* \rangle

lemma *conv-rel-to-sign*:
fixes $f :: 'a :: \text{order} \Rightarrow \text{real}$
shows *monotone* (\leq) $(\leq_{\eta}) f =$ *mono* $((*)(\pm_{\eta}) \circ f)$
 \langle *proof* \rangle

instantiation *RelDirection* :: *times*

begin

definition *times-RelDirection* :: *RelDirection* \Rightarrow *RelDirection* \Rightarrow *RelDirection* **where**
times-RelDirection-def: *times-RelDirection* $x y =$ (if $x = y$ then *Fwd* else *Rev*)

instance \langle *proof* \rangle
end

lemmas *rel-dir-mult[simp]* = *times-RelDirection-def*

lemma *dir-mult-hom*: $(\pm_{\sigma} * \tau) = (\pm_{\sigma}) * ((\pm_{\tau}) :: \text{real})$

<proof>

Additional lemmas about clamp for the specific case on reals.

lemma *clamp-eqI2*:

assumes $x \in \{a..b::real\}$

shows $x = clamp\ a\ b\ x$

<proof>

lemma *clamp-eqI*:

assumes $|x| \leq (a::real)$

shows $x = clamp\ (-a)\ a\ x$

<proof>

lemma *clamp-real-def*:

fixes $x :: real$

shows $clamp\ a\ b\ x = max\ a\ (min\ x\ b)$

<proof>

lemma *clamp-range*:

assumes $a \leq b$

shows $\bigwedge x. clamp\ a\ b\ x \geq a \wedge x. clamp\ a\ b\ x \leq b$ *range* $(clamp\ a\ b) \subseteq \{a..b::real\}$

<proof>

lemma *clamp-abs-le*:

assumes $a \geq (0::real)$

shows $|clamp\ (-a)\ a\ x| \leq |x|$

<proof>

lemma *bounded-clamp*:

fixes $a\ b :: real$

shows *bounded* $((clamp\ a\ b \circ f) ' S)$

<proof>

lemma *bounded-clamp-alt*:

fixes $a\ b :: real$

shows *bounded* $((\lambda x. clamp\ a\ b\ (f\ x)) ' S)$

<proof>

lemma *clamp-borel[measurable]*:

fixes $a\ b :: 'a::\{euclidean-space,second-countable-topology\}$

shows $clamp\ a\ b \in borel\text{-measurable}\ borel$

<proof>

lemma *monotone-clamp*:

assumes *monotone* (\leq) $(\leq_{\eta})\ f$

shows *monotone* (\leq) $(\leq_{\eta})\ (\lambda\omega. clamp\ a\ (b::real)\ (f\ \omega))$

<proof>

This part introduces the term *KL-div* as the Kullback-Leibler divergence

between a pair of Bernoulli random variables. The expression is useful to express some of the Chernoff bounds more concisely [12, Th. 1].

lemma *radon-nikodym-pmf*:

assumes $set\text{-}pmf\ p \subseteq set\text{-}pmf\ q$

defines $f \equiv (\lambda x. ennreal\ (pmf\ p\ x / pmf\ q\ x))$

shows

$AE\ x\ in\ measure\text{-}pmf\ q. RN\text{-}deriv\ q\ p\ x = f\ x\ (is\ ?R1)$

$AE\ x\ in\ measure\text{-}pmf\ p. RN\text{-}deriv\ q\ p\ x = f\ x\ (is\ ?R2)$

$\langle proof \rangle$

lemma *KL-divergence-pmf*:

assumes $set\text{-}pmf\ q \subseteq set\text{-}pmf\ p$

shows $KL\text{-}divergence\ b\ (measure\text{-}pmf\ p)\ (measure\text{-}pmf\ q) = (\int x. log\ b\ (pmf\ q\ x / pmf\ p\ x)\ \partial q)$

$\langle proof \rangle$

definition *KL-div* :: $real \Rightarrow real \Rightarrow real$ **where**

$KL\text{-}div\ p\ q = KL\text{-}divergence\ (exp\ 1)\ (bernoulli\text{-}pmf\ q)\ (bernoulli\text{-}pmf\ p)$

lemma *KL-div-eq*:

assumes $q \in \{0 < .. < 1\}\ p \in \{0..1\}$

shows $KL\text{-}div\ p\ q = p * ln\ (p/q) + (1-p) * ln\ ((1-p)/(1-q))\ (is\ ?L = ?R)$

$\langle proof \rangle$

lemma *KL-div-extreme-cases*:

assumes $p \in \{0,1\}$

shows $KL\text{-}div\ p\ p = 0\ (is\ ?L = ?R)$

$\langle proof \rangle$

lemma *KL-div-eq'*:

assumes $q \in \{0..1\}\ p \in \{0..1\}\ q > 0 \vee p = 0\ q < 1 \vee p = 1$

shows $KL\text{-}div\ p\ q = p * ln\ (p/q) + (1-p) * ln\ ((1-p)/(1-q))\ (is\ ?L = ?R)$

$\langle proof \rangle$

lemma *KL-div-swap-gen*:

assumes $q \in \{0..1\}\ p \in \{0..1\}\ q > 0 \vee p = 0\ q < 1 \vee p = 1$

shows $KL\text{-}div\ p\ q = KL\text{-}div\ (1-p)\ (1-q)$

$\langle proof \rangle$

lemma *KL-div-swap*:

assumes $q \in \{0 < .. < 1\}\ p \in \{0..1\}$

shows $KL\text{-}div\ p\ q = KL\text{-}div\ (1-p)\ (1-q)$

$\langle proof \rangle$

A few results about independent random variables:

lemma (in *prob-space*) *indep-vars-const*:

assumes $\bigwedge i. i \in I \implies c\ i \in space\ (N\ i)$

shows $indep\text{-}vars\ N\ (\lambda i. c\ i)\ I$

$\langle proof \rangle$

lemma *indep-vars-map-pmf*:

assumes *prob-space.indep-vars* (*measure-pmf* *p*) ($\lambda\cdot$. *discrete*) (λi . $X\ i \circ f$) *I*
shows *prob-space.indep-vars* (*map-pmf* *f p*) ($\lambda\cdot$. *discrete*) $X\ I$
<proof>

lemma *indep-var-pair-pmf*:

fixes $x\ y :: 'a\ pmf$
shows *prob-space.indep-var* (*pair-pmf* $x\ y$) *discrete fst discrete snd*
<proof>

lemma *measure-pair-pmf*: $measure\ (pair-pmf\ p\ q)\ (A \times B) = measure\ p\ A * measure\ q\ B$ (**is** $?L = ?R$)
<proof>

instance *bool* :: *second-countable-topology*
<proof>

end

2 Definition

This section introduces the concept of negatively associated random variables (RVs). The definition follows, as closely as possible, the original description by Joag-Dev and Proschan [13].

However, the following modifications have been made:

Singleton and empty sets of random variables are considered negatively associated. This is useful because it simplifies many of the induction proofs. The second modification is that the RV's don't have to be real valued. Instead the range can be into any linearly ordered space with the borel σ -algebra. This is a major enhancement compared to the original work, as well as results by following authors [6, 7, 8, 14, 17].

theory *Negative-Association-Definition*

imports

Concentration-Inequalities.Bienaymes-Identity

Negative-Association-Util

begin

context *prob-space*

begin

definition *neg-assoc* :: $('i \Rightarrow 'a \Rightarrow 'c :: \{linorder-topology\}) \Rightarrow 'i\ set \Rightarrow bool$

where *neg-assoc* $X\ I =$ ($\forall i \in I$. *random-variable borel* ($X\ i$)) \wedge

$(\forall (f :: nat \Rightarrow ('i \Rightarrow 'c) \Rightarrow real)\ J. J \subseteq I \wedge$

$(\forall \iota < 2$. *bounded* (*range* ($f\ \iota$)) \wedge *mono*($f\ \iota$) \wedge *depends-on* ($f\ \iota$) ($[J, I - J]!\iota$)) \wedge

$f \iota \in \text{PiM } ([J, I - J]! \iota) (\lambda \cdot. \text{borel}) \rightarrow_M \text{borel} \longrightarrow$
covariance ($f \ 0 \circ \text{flip } X$) ($f \ 1 \circ \text{flip } X \leq 0$)

lemma *neg-assocI*:

assumes $\bigwedge i. i \in I \implies \text{random-variable borel } (X \ i)$
assumes $\bigwedge f \ g \ J. J \subseteq I$
 $\implies \text{depends-on } f \ J \implies \text{depends-on } g \ (I - J)$
 $\implies \text{mono } f \implies \text{mono } g$
 $\implies \text{bounded } (\text{range } f :: \text{real set}) \implies \text{bounded } (\text{range } g)$
 $\implies f \in \text{PiM } J (\lambda \cdot. \text{borel}) \rightarrow_M \text{borel} \implies g \in \text{PiM } (I - J) (\lambda \cdot. \text{borel}) \rightarrow_M \text{borel}$
 $\implies \text{covariance } (f \circ \text{flip } X) (g \circ \text{flip } X) \leq 0$
shows *neg-assoc* $X \ I$
 $\langle \text{proof} \rangle$

lemma *neg-assocI2*:

assumes [*measurable*]: $\bigwedge i. i \in I \implies \text{random-variable borel } (X \ i)$
assumes $\bigwedge f \ g \ J. J \subseteq I$
 $\implies \text{depends-on } f \ J \implies \text{depends-on } g \ (I - J)$
 $\implies \text{mono } f \implies \text{mono } g$
 $\implies \text{bounded } (\text{range } f) \implies \text{bounded } (\text{range } g)$
 $\implies f \in \text{PiM } J (\lambda \cdot. \text{borel}) \rightarrow_M (\text{borel} :: \text{real measure})$
 $\implies g \in \text{PiM } (I - J) (\lambda \cdot. \text{borel}) \rightarrow_M (\text{borel} :: \text{real measure})$
 $\implies (\int \omega. f(\lambda i. X \ i \ \omega) * g(\lambda i. X \ i \ \omega) \ \partial M) \leq (\int \omega. f(\lambda i. X \ i \ \omega) \ \partial M) * (\int \omega. g(\lambda i. X \ i \ \omega) \ \partial M)$
shows *neg-assoc* $X \ I$
 $\langle \text{proof} \rangle$

lemma *neg-assoc-empty*:

neg-assoc $X \ \{\}$
 $\langle \text{proof} \rangle$

lemma *neg-assoc-singleton*:

assumes *random-variable borel* $(X \ i)$
shows *neg-assoc* $X \ \{i\}$
 $\langle \text{proof} \rangle$

lemma *neg-assoc-imp-measurable*:

assumes *neg-assoc* $X \ I$
assumes $i \in I$
shows *random-variable borel* $(X \ i)$
 $\langle \text{proof} \rangle$

Even though the assumption was that defining property is true for pairs of monotone functions over the random variables, it is also true for pairs of anti-monotone functions.

lemma *neg-assoc-imp-mult-mono-bounded*:

fixes $f \ g :: ('i \Rightarrow 'c :: \text{linorder-topology}) \Rightarrow \text{real}$
assumes *neg-assoc* $X \ I$
assumes $J \subseteq I$

assumes *bounded* (*range f*) *bounded* (*range g*)
assumes *monotone* (\leq) ($\leq_{\geq\eta}$) *f monotone* (\leq) ($\leq_{\geq\eta}$) *g*
assumes *depends-on f J depends-on g* ($I-J$)
assumes [*measurable*]: *f* \in *borel-measurable* ($Pi_M J (\lambda-. \text{borel})$)
assumes [*measurable*]: *g* \in *borel-measurable* ($Pi_M (I-J) (\lambda-. \text{borel})$)
shows
covariance ($f \circ \text{flip } X$) ($g \circ \text{flip } X$) ≤ 0
 $(\int \omega. f (\lambda i. X i \omega) * g (\lambda i. X i \omega) \partial M) \leq \text{expectation } (\lambda x. f(\lambda y. X y x)) * \text{expectation } (\lambda x. g(\lambda y. X y x))$
(is ?L \leq ?R)
 $\langle \text{proof} \rangle$

lemma *lim-min-n*: $(\lambda n. \text{min } (\text{real } n) x) \longrightarrow x$
 $\langle \text{proof} \rangle$

lemma *lim-clamp-n*: $(\lambda n. \text{clamp } (-\text{real } n) (\text{real } n) x) \longrightarrow x$
 $\langle \text{proof} \rangle$

lemma *neg-assoc-imp-mult-mono*:
fixes *f g* :: ($'i \Rightarrow 'c::\text{linorder-topology}$) \Rightarrow *real*
assumes *neg-assoc X I*
assumes $J \subseteq I$
assumes *square-integrable M* ($f \circ \text{flip } X$) *square-integrable M* ($g \circ \text{flip } X$)
assumes *monotone* (\leq) ($\leq_{\geq\eta}$) *f monotone* (\leq) ($\leq_{\geq\eta}$) *g*
assumes *depends-on f J depends-on g* ($I-J$)
assumes [*measurable*]: *f* \in *borel-measurable* ($Pi_M J (\lambda-. \text{borel})$)
assumes [*measurable*]: *g* \in *borel-measurable* ($Pi_M (I-J) (\lambda-. \text{borel})$)
shows $(\int \omega. f (\lambda i. X i \omega) * g (\lambda i. X i \omega) \partial M) \leq (\int x. f(\lambda y. X y x) \partial M) * (\int x. g(\lambda y. X y x) \partial M)$
(is ?L \leq ?R)
 $\langle \text{proof} \rangle$

Property P4 [13]

lemma *neg-assoc-subset*:
assumes $J \subseteq I$
assumes *neg-assoc X I*
shows *neg-assoc X J*
 $\langle \text{proof} \rangle$

lemma *neg-assoc-imp-mult-mono-nonneg*:
fixes *f g* :: ($'i \Rightarrow 'c::\text{linorder-topology}$) \Rightarrow *real*
assumes *neg-assoc X I J* $J \subseteq I$
assumes *range f* $\subseteq \{0..\}$ *range g* $\subseteq \{0..\}$
assumes *integrable M* ($f \circ \text{flip } X$) *integrable M* ($g \circ \text{flip } X$)
assumes *monotone* (\leq) ($\leq_{\geq\eta}$) *f monotone* (\leq) ($\leq_{\geq\eta}$) *g*
assumes *depends-on f J depends-on g* ($I-J$)
assumes *f* \in *borel-measurable* ($Pi_M J (\lambda-. \text{borel})$) *g* \in *borel-measurable* ($Pi_M (I-J) (\lambda-. \text{borel})$)
shows *has-int-that M* ($\lambda \omega. f (\text{flip } X \omega) * g (\text{flip } X \omega)$)

($\lambda r. r \leq \text{expectation } (f \circ \text{flip } X) * \text{expectation } (g \circ \text{flip } X)$)
 <proof>

Property P2 [13]

lemma *neg-assoc-imp-prod-mono*:

fixes $f :: 'i \Rightarrow ('c::\text{linorder-topology}) \Rightarrow \text{real}$
assumes *finite* I
assumes *neg-assoc* $X I$
assumes $\bigwedge i. i \in I \implies \text{integrable } M (\lambda \omega. f i (X i \omega))$
assumes $\bigwedge i. i \in I \implies \text{monotone } (\leq) (\leq_{\eta}) (f i)$
assumes $\bigwedge i. i \in I \implies \text{range } (f i) \subseteq \{0..\}$
assumes $\bigwedge i. i \in I \implies f i \in \text{borel-measurable borel}$
shows *has-int-that* $M (\lambda \omega. (\prod_{i \in I}. f i (X i \omega))) (\lambda r. r \leq (\prod_{i \in I}. \text{expectation } (\lambda \omega. f i (X i \omega))))$
 <proof>

Property P5 [13]

lemma *neg-assoc-compose*:

fixes $f :: 'j \Rightarrow ('i \Rightarrow ('c::\text{linorder-topology})) \Rightarrow ('d :: \text{linorder-topology})$
assumes *finite* I
assumes *neg-assoc* $X I$
assumes $\bigwedge j. j \in J \implies \text{deps } j \subseteq I$
assumes $\bigwedge j1 j2. j1 \in J \implies j2 \in J \implies j1 \neq j2 \implies \text{deps } j1 \cap \text{deps } j2 = \{\}$
assumes $\bigwedge j. j \in J \implies \text{monotone } (\leq) (\leq_{\eta}) (f j)$
assumes $\bigwedge j. j \in J \implies \text{depends-on } (f j) (\text{deps } j)$
assumes $\bigwedge j. j \in J \implies f j \in \text{borel-measurable } (PiM (\text{deps } j) (\lambda \cdot. \text{borel}))$
shows *neg-assoc* $(\lambda j \omega. f j (\lambda i. X i \omega)) J$
 <proof>

lemma *neg-assoc-compose-simple*:

fixes $f :: 'i \Rightarrow ('c::\text{linorder-topology}) \Rightarrow ('d :: \text{linorder-topology})$
assumes *finite* I
assumes *neg-assoc* $X I$
assumes $\bigwedge i. i \in I \implies \text{monotone } (\leq) (\leq_{\eta}) (f i)$
assumes [*measurable*]: $\bigwedge i. i \in I \implies f i \in \text{borel-measurable borel}$
shows *neg-assoc* $(\lambda i \omega. f i (X i \omega)) I$
 <proof>

lemma *covariance-distr*:

fixes $f g :: 'b \Rightarrow \text{real}$
assumes [*measurable*]: $\varphi \in M \rightarrow_M N$ $f \in \text{borel-measurable } N$ $g \in \text{borel-measurable } N$
shows *prob-space.covariance* $(\text{distr } M N \varphi) f g = \text{covariance } (f \circ \varphi) (g \circ \varphi)$ (**is** $?L = ?R$)
 <proof>

lemma *neg-assoc-iff-distr*:

assumes [*measurable*]: $\bigwedge i. i \in I \implies X i \in \text{borel-measurable } M$
shows *neg-assoc* $X I \longleftrightarrow$

$prob\text{-}space.neg\text{-}assoc (distr M (PiM I (\lambda\cdot. borel))) (\lambda\omega. \lambda i \in I. X i \omega) (flip id) I$
 $(is ?L \longleftrightarrow ?R)$
 <proof>

lemma *neg-assoc-cong*:

assumes *finite I*
assumes [*measurable*]: $\bigwedge i. i \in I \implies Y i \in borel\text{-}measurable M$
assumes *neg-assoc X I* $\bigwedge i. i \in I \implies AE \omega \text{ in } M. X i \omega = Y i \omega$
shows *neg-assoc Y I*
 <proof>

lemma *neg-assoc-reindex-aux*:

assumes *inj-on h I*
assumes *neg-assoc X (h ' I)*
shows *neg-assoc* $(\lambda k. X (h k)) I$
 <proof>

lemma *neg-assoc-reindex*:

assumes *inj-on h I finite I*
shows *neg-assoc X (h ' I)* \longleftrightarrow *neg-assoc* $(\lambda k. X (h k)) I$ (**is** *?L* \longleftrightarrow *?R*)
 <proof>

lemma *measurable-compose-merge-1*:

assumes *depends-on h K*
assumes $h \in PiM K M' \rightarrow_M N K \subseteq I \cup J$
assumes $(\lambda x. restrict (fst (f x)) (K \cap I)) \in A \rightarrow_M PiM (K \cap I) M'$
assumes $(\lambda x. restrict (snd (f x)) (K \cap J)) \in A \rightarrow_M PiM (K \cap J) M'$
shows $(\lambda x. h(merge I J (f x))) \in A \rightarrow_M N$
 <proof>

lemma *measurable-compose-merge-2*:

assumes *depends-on h K h* $h \in PiM K M' \rightarrow_M N K \subseteq I \cup J$
assumes $(\lambda x. restrict (f x) (K \cap I)) \in A \rightarrow_M PiM (K \cap I) M'$
assumes $(\lambda x. restrict (g x) (K \cap J)) \in A \rightarrow_M PiM (K \cap J) M'$
shows $(\lambda x. h(merge I J (f x, g x))) \in A \rightarrow_M N$
 <proof>

lemma *neg-assoc-combine*:

fixes $I I1 I2 :: 'i \text{ set}$
fixes $X :: 'i \Rightarrow 'a \Rightarrow ('b::linorder\text{-}topology)$
assumes *finite I I1 I2* $= I I1 \cap I2 = \{\}$
assumes *indep-var* $(PiM I1 (\lambda\cdot. borel)) (\lambda\omega. \lambda i \in I1. X i \omega) (PiM I2 (\lambda\cdot. borel))$
 $(\lambda\omega. \lambda i \in I2. X i \omega)$
assumes *neg-assoc X I1*
assumes *neg-assoc X I2*
shows *neg-assoc X I*
 <proof>

Property P7 [13]

```

lemma neg-assoc-union:
  fixes  $I :: 'i$  set
  fixes  $p :: 'j \Rightarrow 'i$  set
  fixes  $X :: 'i \Rightarrow 'a \Rightarrow ('b::linorder-topology)$ 
  assumes finite  $I \cup (p \text{ ` } J) = I$ 
  assumes indep-vars  $(\lambda j. \text{PiM } (p \ j) (\lambda \cdot. \text{borel})) (\lambda j \ \omega. \lambda i \in p \ j. X \ i \ \omega) \ J$ 
  assumes  $\bigwedge j. j \in J \implies \text{neg-assoc } X \ (p \ j)$ 
  assumes disjoint-family-on  $p \ J$ 
  shows neg-assoc  $X \ I$ 
<proof>

```

Property P5 [13]

```

lemma indep-imp-neg-assoc:
  assumes finite  $I$ 
  assumes indep-vars  $(\lambda \cdot. \text{borel}) \ X \ I$ 
  shows neg-assoc  $X \ I$ 
<proof>

```

end

```

lemma neg-assoc-map-pmf:
  shows measure-pmf.neg-assoc  $(\text{map-pmf } f \ p) \ X \ I = \text{measure-pmf.neg-assoc } p \ (\lambda i$ 
 $\ \omega. X \ i \ (f \ \omega)) \ I$ 
  (is  $?L \longleftrightarrow ?R$ )
<proof>

```

end

3 Chernoff-Hoeffding Bounds

This section shows that all the well-known Chernoff-Hoeffding bounds hold also for negatively associated random variables. The proofs follow the derivations by Hoeffding [11], as well as, Motwani and Raghavan [16, Ch. 4], with the modification that the crucial steps, where the classic proofs use independence, are replaced with the application of Property P2 for negatively associated RV's.

```

theory Negative-Association-Chernoff-Bounds
  imports
    Negative-Association-Definition
    Concentration-Inequalities.McDiarmid-Inequality
    Weighted-Arithmetic-Geometric-Mean.Weighted-Arithmetic-Geometric-Mean
begin

context prob-space
begin

context

```

fixes $I :: 'i \text{ set}$
fixes $X :: 'i \Rightarrow 'a \Rightarrow \text{real}$
assumes $na\text{-}X$: $\text{neg-assoc } X \ I$
assumes $fin\text{-}I$: $\text{finite } I$
begin

private lemma *transfer-to-clamped-vars*:

assumes $(\forall i \in I. AE \ \omega \ \text{in } M. X \ i \ \omega \in \{a \ i..b \ i\} \wedge a \ i \leq b \ i)$
assumes $\mathcal{X}\text{-def}$: $\mathcal{X} = (\lambda i. \text{clamp } (a \ i) \ (b \ i) \circ X \ i)$
shows $\text{neg-assoc } \mathcal{X} \ I$ (**is** $?A$)
and $\bigwedge i. i \in I \implies \text{expectation } (\mathcal{X} \ i) = \text{expectation } (X \ i)$
and $\mathcal{P}(\omega \ \text{in } M. (\sum i \in I. X \ i \ \omega) \leq_{\geq \eta} c) = \mathcal{P}(\omega \ \text{in } M. (\sum i \in I. \mathcal{X} \ i \ \omega) \leq_{\geq \eta} c)$ (**is** $?C$)
and $\bigwedge i \ \omega. i \in I \implies \mathcal{X} \ i \ \omega \in \{a \ i..b \ i\}$
and $\bigwedge i \ S. i \in I \implies \text{bounded } (\mathcal{X} \ i \ 'S)$
and $\bigwedge i. i \in I \implies \text{expectation } (\mathcal{X} \ i) \in \{a \ i..b \ i\}$
 $\langle \text{proof} \rangle$

lemma *ln-one-plus-x-lower-bound*:

assumes $x \geq (0::\text{real})$
shows $2*x/(2+x) \leq \ln(1+x)$
 $\langle \text{proof} \rangle$

Based on Theorem 4.1 by Motwani and Raghavan [16].

theorem *multiplicative-chernoff-bound-upper*:

assumes $\delta > 0$
assumes $\bigwedge i. i \in I \implies AE \ \omega \ \text{in } M. X \ i \ \omega \in \{0..1\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
shows $\mathcal{P}(\omega \ \text{in } M. (\sum i \in I. X \ i \ \omega) \geq (1+\delta) * \mu) \leq (\exp \ \delta / ((1+\delta) \ \text{powr } (1+\delta)))$
 $\text{powr } \mu$ (**is** $?L \leq ?R$)
and $\mathcal{P}(\omega \ \text{in } M. (\sum i \in I. X \ i \ \omega) \geq (1+\delta) * \mu) \leq \exp(-(\delta^2) * \mu / (2+\delta))$
(is $\leq ?R1$)
 $\langle \text{proof} \rangle$

lemma *ln-one-minus-x-lower-bound*:

assumes $x \in \{(0::\text{real})..<1\}$
shows $(x^2/2-x)/(1-x) \leq \ln(1-x)$
 $\langle \text{proof} \rangle$

Based on Theorem 4.2 by Motwani and Raghavan [16].

theorem *multiplicative-chernoff-bound-lower*:

assumes $\delta \in \{0<..<1\}$
assumes $\bigwedge i. i \in I \implies AE \ \omega \ \text{in } M. X \ i \ \omega \in \{0..1\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
shows $\mathcal{P}(\omega \ \text{in } M. (\sum i \in I. X \ i \ \omega) \leq (1-\delta)*\mu) \leq (\exp(-\delta)/(1-\delta) \ \text{powr } (1-\delta))$
 $\text{powr } \mu$ (**is** $?L \leq ?R$)
and $\mathcal{P}(\omega \ \text{in } M. (\sum i \in I. X \ i \ \omega) \leq (1-\delta)*\mu) \leq (\exp(-(\delta^2)*\mu/2))$ (**is** $\leq ?R1$)
 $\langle \text{proof} \rangle$

theorem *multiplicative-chernoff-bound-two-sided:*

assumes $\delta \in \{0 < \dots < 1\}$
assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{0..1\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
shows $\mathcal{P}(\omega \text{ in } M. |\sum i \in I. X \ i \ \omega - \mu| \geq \delta * \mu) \leq 2 * (\exp(-(\delta^2) * \mu / 3))$ **(is**
 $?L \leq ?R)$
 $\langle \text{proof} \rangle$

lemma *additive-chernoff-bound-upper-aux:*

assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{0..1\} \ I \neq \{\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i)) / \text{real } (\text{card } I)$
assumes $\delta \in \{0 < \dots < 1 - \mu\} \ \mu \in \{0 < \dots < 1\}$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. X \ i \ \omega) \geq (\mu + \delta) * \text{real } (\text{card } I)) \leq \exp(-\text{real } (\text{card } I) * KL\text{-div}$
 $* KL\text{-div } (\mu + \delta) \ \mu)$
(is $?L \leq ?R)$
 $\langle \text{proof} \rangle$

lemma *additive-chernoff-bound-upper-aux-2:*

assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{0..1\} \ I \neq \{\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i)) / \text{real } (\text{card } I)$
assumes $\mu \in \{0 < \dots < 1\}$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. X \ i \ \omega) \geq \text{real } (\text{card } I)) \leq \exp(-\text{real } (\text{card } I) * KL\text{-div}$
 $1 \ \mu)$
(is $?L \leq ?R)$
 $\langle \text{proof} \rangle$

Based on Theorem 1 by Hoeffding [11].

lemma *additive-chernoff-bound-upper:*

assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{0..1\} \ I \neq \{\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i)) / \text{real } (\text{card } I)$
assumes $\delta \in \{0..1 - \mu\} \ \mu \in \{0 < \dots < 1\}$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. X \ i \ \omega) \geq (\mu + \delta) * \text{real } (\text{card } I)) \leq \exp(-\text{real } (\text{card } I) * KL\text{-div}$
 $* KL\text{-div } (\mu + \delta) \ \mu)$
(is $?L \leq ?R)$
 $\langle \text{proof} \rangle$

Based on Theorem 2 by Hoeffding [11].

lemma *hoeffding-bound-upper:*

assumes $\bigwedge i. i \in I \implies a \ i \leq b \ i$
assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{a \ i..b \ i\}$
defines $n \equiv \text{real } (\text{card } I)$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
assumes $\delta \geq 0 \ (\sum i \in I. (b \ i - a \ i)^2) > 0$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. X \ i \ \omega) \geq \mu + \delta * n) \leq \exp(-2 * (n * \delta)^2 / (\sum i \in I.$
 $(b \ i - a \ i)^2))$
(is $?L \leq ?R)$
 $\langle \text{proof} \rangle$

end

Dual and two-sided versions of Theorem 1 and 2 by Hoeffding [11].

lemma *additive-chernoff-bound-lower*:

assumes *neg-assoc* X I *finite* I
assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{0..1\} \ I \neq \{\}$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i)) / \text{real } (\text{card } I)$
assumes $\delta \in \{0.. \mu\} \ \mu \in \{0 < .. < I\}$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. \ X \ i \ \omega) \leq (\mu - \delta) * \text{real } (\text{card } I)) \leq \text{exp } (-\text{real } (\text{card } I) * \text{KL-div } (\mu - \delta) \ \mu)$
(is ?L ≤ ?R)
<proof>

lemma *hoeffding-bound-lower*:

assumes *neg-assoc* X I *finite* I
assumes $\bigwedge i. i \in I \implies a \ i \leq b \ i$
assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{a \ i..b \ i\}$
defines $n \equiv \text{real } (\text{card } I)$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
assumes $\delta \geq 0 \ (\sum i \in I. (b \ i - a \ i)^{\wedge 2}) > 0$
shows $\mathcal{P}(\omega \text{ in } M. (\sum i \in I. \ X \ i \ \omega) \leq \mu - \delta * n) \leq \text{exp } (-2 * (n * \delta)^{\wedge 2} / (\sum i \in I. (b \ i - a \ i)^{\wedge 2}))$
(is ?L ≤ ?R)
<proof>

lemma *hoeffding-bound-two-sided*:

assumes *neg-assoc* X I *finite* I
assumes $\bigwedge i. i \in I \implies a \ i \leq b \ i$
assumes $\bigwedge i. i \in I \implies AE \ \omega \text{ in } M. \ X \ i \ \omega \in \{a \ i..b \ i\} \ I \neq \{\}$
defines $n \equiv \text{real } (\text{card } I)$
defines $\mu \equiv (\sum i \in I. \text{expectation } (X \ i))$
assumes $\delta \geq 0 \ (\sum i \in I. (b \ i - a \ i)^{\wedge 2}) > 0$
shows $\mathcal{P}(\omega \text{ in } M. |(\sum i \in I. \ X \ i \ \omega) - \mu| \geq \delta * n) \leq 2 * \text{exp } (-2 * (n * \delta)^{\wedge 2} / (\sum i \in I. (b \ i - a \ i)^{\wedge 2}))$
(is ?L ≤ ?R)
<proof>

end

end

4 The FKG inequality

The FKG inequality [9] is a generalization of Chebyshev's less known other inequality. It is sometimes referred to as Chebyshev's sum inequality. Although there is also a continuous version, which can be stated as:

$$E[fg] \geq E[f]E[g]$$

where f, g are continuous simultaneously monotone or simultaneously antimonotone functions on the Lebesgue probability space $[a, b] \subseteq \mathbb{R}$. (Ef denotes the expectation of the function.)

Note that the inequality is also true for totally ordered discrete probability spaces, for example: $\{1, \dots, n\}$ with uniform probabilities.

The FKG inequality is essentially a generalization of the above to not necessarily totally ordered spaces, but finite distributive lattices.

The proof follows the derivation in the book by Alon and Spencer [2, Ch. 6].

theory *Negative-Association-FKG-Inequality*

imports

Negative-Association-Util

Birkhoff-Finite-Distributive-Lattices.Birkhoff-Finite-Distributive-Lattices

begin

theorem *four-functions-helper:*

fixes $\varphi :: \text{nat} \Rightarrow 'a \text{ set} \Rightarrow \text{real}$

assumes *finite I*

assumes $\bigwedge i. i \in \{0..3\} \implies \varphi i \in \text{Pow } I \rightarrow \{0..\}$

assumes $\bigwedge A B. A \subseteq I \implies B \subseteq I \implies \varphi 0 A * \varphi 1 B \leq \varphi 2 (A \cup B) * \varphi 3 (A \cap B)$

shows $(\sum A \in \text{Pow } I. \varphi 0 A) * (\sum B \in \text{Pow } I. \varphi 1 B) \leq (\sum C \in \text{Pow } I. \varphi 2 C) * (\sum D \in \text{Pow } I. \varphi 3 D)$

<proof>

The following is the Ahlswede-Daykin inequality [1] also referred to by Alon and Spencer as the four functions theorem [2, Th. 6.1.1].

theorem *four-functions:*

fixes $\alpha \beta \gamma \delta :: 'a \text{ set} \Rightarrow \text{real}$

assumes *finite I*

assumes $\alpha \in \text{Pow } I \rightarrow \{0..\} \beta \in \text{Pow } I \rightarrow \{0..\} \gamma \in \text{Pow } I \rightarrow \{0..\} \delta \in \text{Pow } I \rightarrow \{0..\}$

assumes $\bigwedge A B. A \subseteq I \implies B \subseteq I \implies \alpha A * \beta B \leq \gamma (A \cup B) * \delta (A \cap B)$

assumes $M \subseteq \text{Pow } I \ N \subseteq \text{Pow } I$

shows $(\sum A \in M. \alpha A) * (\sum B \in N. \beta B) \leq (\sum C | \exists A \in M. \exists B \in N. C = A \cup B. \gamma C) * (\sum D | \exists A \in M. \exists B \in N. D = A \cap B. \delta D)$

(is ?L ≤ ?R)

<proof>

Using Birkhoff's Representation Theorem [3, 5] it is possible to generalize the previous to finite distributive lattices [2, Cor. 6.1.2].

lemma *four-functions-in-lattice:*

fixes $\alpha \beta \gamma \delta :: 'a :: \text{finite-distrib-lattice} \Rightarrow \text{real}$

assumes $\text{range } \alpha \subseteq \{0..\} \text{ range } \beta \subseteq \{0..\} \text{ range } \gamma \subseteq \{0..\} \text{ range } \delta \subseteq \{0..\}$

assumes $\bigwedge x y. \alpha x * \beta y \leq \gamma (x \sqcup y) * \delta (x \sqcap y)$

shows $(\sum x \in M. \alpha x) * (\sum y \in N. \beta y) \leq (\sum c | \exists x \in M. \exists y \in N. c = x \sqcup y. \gamma c) * (\sum d | \exists x \in M. \exists y \in N. d = x \sqcap y. \delta d)$

(is ?L ≤ ?R)
 ⟨proof⟩

theorem *fkf-inequality*:

fixes $\mu :: 'a :: \text{finite-distrib-lattice} \Rightarrow \text{real}$
assumes $\text{range } \mu \subseteq \{0..\}$ $\text{range } f \subseteq \{0..\}$ $\text{range } g \subseteq \{0..\}$
assumes $\bigwedge x y. \mu x * \mu y \leq \mu (x \sqcup y) * \mu (x \sqcap y)$
assumes *mono f mono g*
shows $(\sum x \in \text{UNIV}. \mu x * f x) * (\sum x \in \text{UNIV}. \mu x * g x) \leq (\sum x \in \text{UNIV}. \mu x * f x * g x) * \text{sum } \mu \text{ UNIV}$
 (is ?L ≤ ?R)
 ⟨proof⟩

theorem *fkf-inequality-gen*:

fixes $\mu :: 'a :: \text{finite-distrib-lattice} \Rightarrow \text{real}$
assumes $\text{range } \mu \subseteq \{0..\}$
assumes $\bigwedge x y. \mu x * \mu y \leq \mu (x \sqcup y) * \mu (x \sqcap y)$
assumes *monotone (≤) (≤≥_τ) f monotone (≤) (≤≥_σ) g*
shows $(\sum x \in \text{UNIV}. \mu x * f x) * (\sum x \in \text{UNIV}. \mu x * g x) \leq_{\tau * \sigma} (\sum x \in \text{UNIV}. \mu x * f x * g x) * \text{sum } \mu \text{ UNIV}$
 (is ?L ≤≥_{?x} ?R)
 ⟨proof⟩

theorem *fkf-inequality-pmf*:

fixes $M :: ('a :: \text{finite-distrib-lattice}) \text{ pmf}$
fixes $f g :: 'a \Rightarrow \text{real}$
assumes $\bigwedge x y. \text{pmf } M x * \text{pmf } M y \leq \text{pmf } M (x \sqcup y) * \text{pmf } M (x \sqcap y)$
assumes *monotone (≤) (≤≥_τ) f monotone (≤) (≤≥_σ) g*
shows $(\int x. f x \partial M) * (\int x. g x \partial M) \leq_{\tau * \sigma} (\int x. f x * g x \partial M)$
 (is ?L ≤≥₋ ?R)
 ⟨proof⟩

end

5 Preliminary Results on Lattices

This entry establishes a few missing lemmas for the set-based theory of lattices from “HOL-Algebra”. In particular, it introduces the sublocale for distributive lattices.

More crucially, a transfer theorem which can be used in conjunction with the Types-To-Sets mechanism to be able to work with locally defined finite distributive lattices.

This is being needed for the verification of the negative association of permutation distributions in Section 6.

theory *Negative-Association-More-Lattices*

imports *HOL-Algebra.Lattice*

begin

Lemma 1 Birkhoff Lattice Theory, p.8, L3

lemma (in *lattice*) *meet-assoc-law*:

assumes $x \in \text{carrier } L$ $y \in \text{carrier } L$ $z \in \text{carrier } L$

shows $x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z$

<proof>

Lemma 1 Birkhoff Lattice Theory, p.8, L3

lemma (in *lattice*) *join-assoc-law*:

assumes $x \in \text{carrier } L$ $y \in \text{carrier } L$ $z \in \text{carrier } L$

shows $x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$

<proof>

Lemma 1 Birkhoff Lattice Theory, p.8, L4

lemma (in *lattice*) *absorbtion-law*:

assumes $x \in \text{carrier } L$ $y \in \text{carrier } L$

shows $x \sqcap (x \sqcup y) = x$ $x \sqcup (x \sqcap y) = x$

<proof>

Theorem 9 Birkhoff Lattice Theory, p.11

lemma (in *lattice*) *distrib-laws-equiv*:

defines *meet-distrib* $\equiv (\forall x y z. \{x,y,z\} \subseteq \text{carrier } L \longrightarrow (x \sqcap (y \sqcup z)) = (x \sqcap y) \sqcup (x \sqcap z))$

defines *join-distrib* $\equiv (\forall x y z. \{x,y,z\} \subseteq \text{carrier } L \longrightarrow (x \sqcup (y \sqcap z)) = (x \sqcup y) \sqcap (x \sqcup z))$

shows *meet-distrib* \longleftrightarrow *join-distrib*

<proof>

lemma (in *lattice*) *lub-unique-set*:

assumes *is-lub* L z S

shows $z = \bigsqcup S$

<proof>

lemma (in *lattice*) *lub-unique*:

assumes *is-lub* L z $\{x,y\}$

shows $z = x \sqcup y$

<proof>

lemma (in *lattice*) *glb-unique-set*:

assumes *is-glb* L z S

shows $z = \bigsqcap S$

<proof>

lemma (in *lattice*) *glb-unique*:

assumes *is-glb* L z $\{x,y\}$

shows $z = x \sqcap y$

<proof>

lemma (in *lattice*) *inf-lower*:

assumes $S \subseteq \text{carrier } L$ $s \in S$ *finite* S
shows $\prod S \sqsubseteq s$
 $\langle \text{proof} \rangle$

lemma (*in lattice*) *sup-upper*:
assumes $S \subseteq \text{carrier } L$ $s \in S$ *finite* S
shows $s \sqsubseteq \bigsqcup S$
 $\langle \text{proof} \rangle$

locale *distrib-lattice = lattice +*
assumes *max-distrib*:
 $x \in \text{carrier } L \implies y \in \text{carrier } L \implies z \in \text{carrier } L \implies (x \sqcap (y \sqcup z)) = (x \sqcap y) \sqcup (x \sqcap z)$
begin

lemma *min-distrib*:
assumes $x \in \text{carrier } L$ $y \in \text{carrier } L$ $z \in \text{carrier } L$
shows $(x \sqcup (y \sqcap z)) = (x \sqcup y) \sqcap (x \sqcup z)$
 $\langle \text{proof} \rangle$

end

locale *finite-ne-distrib-lattice = distrib-lattice +*
assumes *non-empty-carrier*: $\text{carrier } L \neq \{\}$
assumes *finite-carrier*: *finite* ($\text{carrier } L$)
begin

lemma *bounded-lattice-axioms-1*: $\exists x. \text{least } L x (\text{carrier } L)$
 $\langle \text{proof} \rangle$

lemma *bounded-lattice-axioms-2*: $\exists x. \text{greatest } L x (\text{carrier } L)$
 $\langle \text{proof} \rangle$

sublocale *bounded-lattice*
 $\langle \text{proof} \rangle$

lemma *inf-empty*: $\prod \{\} = \top$
 $\langle \text{proof} \rangle$

lemma *inf-closed*: $S \subseteq \text{carrier } L \implies \prod S \in \text{carrier } L$
 $\langle \text{proof} \rangle$

lemma *inf-insert*:
assumes $x \in \text{carrier } L$ $S \subseteq \text{carrier } L$
shows $\prod (\text{insert } x S) = x \sqcap (\prod S)$
 $\langle \text{proof} \rangle$

lemma *sup-empty*: $\bigsqcup \{\} = \perp$
 $\langle \text{proof} \rangle$

lemma *sup-closed*: $S \subseteq \text{carrier } L \implies \bigsqcup S \in \text{carrier } L$
 ⟨proof⟩

lemma *sup-insert*:
assumes $x \in \text{carrier } L$ $S \subseteq \text{carrier } L$
shows $\bigsqcup (\text{insert } x S) = x \sqcup (\bigsqcup S)$
 ⟨proof⟩

lemma *inf-carrier*: $\bigsqcap (\text{carrier } L) = \perp$
 ⟨proof⟩

lemma *sup-carrier*: $\bigsqcup (\text{carrier } L) = \top$
 ⟨proof⟩

lemma *transfer-to-type*:
assumes *finite* (*carrier* L) *type-definition* $\text{Rep } \text{Abs } (\text{carrier } L)$
defines $\text{inf}' \equiv (\lambda M. \text{Abs } (\bigsqcap \text{Rep } 'M))$
defines $\text{sup}' \equiv (\lambda M. \text{Abs } (\bigsqcup \text{Rep } 'M))$
defines $\text{join}' \equiv (\lambda x y. \text{Abs } (\text{Rep } x \sqcap \text{Rep } y))$
defines $\text{le}' \equiv (\lambda x y. (\text{Rep } x \sqsubseteq \text{Rep } y))$
defines $\text{less}' \equiv (\lambda x y. (\text{Rep } x \sqsubset \text{Rep } y))$
defines $\text{meet}' \equiv (\lambda x y. (\text{Abs } (\text{Rep } x \sqcup \text{Rep } y)))$
defines $\text{bot}' \equiv (\text{Abs } \perp :: 'c)$
defines $\text{top}' \equiv \text{Abs } \top$
shows *class.finite-distrib-lattice* $\text{inf}' \text{ sup}' \text{ join}' \text{ le}' \text{ less}' \text{ meet}' \text{ bot}' \text{ top}'$
 ⟨proof⟩

end

end

6 Permutation Distributions

One of the fundamental examples for negatively associated random variables are permutation distributions.

Let x_1, \dots, x_n be n (not-necessarily) distinct values from a totally ordered set, then we choose a permutation $\sigma : \{0, \dots, n-1\} \rightarrow \{0, \dots, n-1\}$ uniformly at random. Then the random variables defined by $X_i(\sigma) = x_{\sigma(i)}$ are negatively associated.

An important special case is the case where x consists of 1 one and $(n-1)$ zeros, modelling randomly putting a ball into one of n bins. Of course the process can be repeated independently, the resulting distribution is also referred to as the balls into bins process. Because of the closure properties established before, it is possible to conclude that the number of hits of each bin in such a process are also negatively associated random variables.

In this section, we will derive that permutation distributions are negatively associated. The proof follows Dubhashi [8, Th. 10] closely. A very short proof was presented in the work by Joag-Dev [13], which, however, is incomplete.

theory *Negative-Association-Permutation-Distributions*

imports

Negative-Association-Definition

Negative-Association-FKG-Inequality

Negative-Association-More-Lattices

Finite-Fields.Finite-Fields-More-PMF

HOL-Types-To-Sets.Types-To-Sets

Executable-Randomized-Algorithms.Randomized-Algorithm

Twelvefold-Way.Card-Bijections

begin

The following introduces a lattice for n -element subsets of a finite set (with size larger or equal to n .) A subset x is smaller or equal to y , if the smallest element of x is smaller or equal to the smallest element of y , the second smallest element of x is smaller or equal to the second smallest element of y , etc.)

The lattice is introduced without name by Dubhashi [7, Example 7].

definition *le-ordered-set-lattice* :: ('a::linorder) set \Rightarrow 'a set \Rightarrow bool

where *le-ordered-set-lattice* $S T = \text{list-all2 } (\leq) (\text{sorted-list-of-set } S) (\text{sorted-list-of-set } T)$

definition *ordered-set-lattice* :: ('a :: linorder) set \Rightarrow nat \Rightarrow 'a set *gorder*

where *ordered-set-lattice* $S n =$

\langle carrier = $\{T. T \subseteq S \wedge \text{finite } T \wedge \text{card } T = n\}$,

eq = (=),

le = *le-ordered-set-lattice* \rangle

definition *osl-repr* :: ('a :: linorder) set \Rightarrow nat \Rightarrow 'a set \Rightarrow nat \Rightarrow 'a

where *osl-repr* $S n e = (\lambda i \in \{..<n\}. \text{sorted-list-of-set } e ! i)$

lemma *osl-carr-sorted-list-of-set*:

assumes *finite* $S n \leq \text{card } S$

assumes $s \in \text{carrier } (\text{ordered-set-lattice } S n)$

defines $t \equiv \text{sorted-list-of-set } s$

shows *finite* $s \text{ card } s = n \ s \subseteq S \ \text{length } t = n \ \text{set } t = s \ \text{sorted-wrt } (<) \ t$

$\langle \text{proof} \rangle$

lemma *ordered-set-lattice-carrier-intro*:

assumes *finite* $S n \leq \text{card } S$

assumes $\text{set } s \subseteq S \ \text{distinct } s \ \text{length } s = n$

shows $\text{set } s \in \text{carrier } (\text{ordered-set-lattice } S n)$

$\langle \text{proof} \rangle$

lemma *osl-list-repr-inj*:

assumes $finite\ S\ n \leq card\ S$
assumes $s \in carrier\ (ordered-set-lattice\ S\ n)$
assumes $t \in carrier\ (ordered-set-lattice\ S\ n)$
assumes $\bigwedge i. osl-repr\ S\ n\ s\ i = osl-repr\ S\ n\ t\ i$
shows $s = t$
 $\langle proof \rangle$

lemma *osl-leD*:

assumes $finite\ S\ n \leq card\ S$
assumes $e \in carrier\ (ordered-set-lattice\ S\ n)$
assumes $f \in carrier\ (ordered-set-lattice\ S\ n)$
shows $e \sqsubseteq_{ordered-set-lattice\ S\ n} f \iff (\forall i. osl-repr\ S\ n\ e\ i \leq osl-repr\ S\ n\ f\ i)$ (**is**
 $?L = ?R$)
 $\langle proof \rangle$

lemma *ordered-set-lattice-partial-order*:

fixes $S :: ('a :: linorder)\ set$
assumes $finite\ S\ n \leq card\ S$
shows *partial-order* $(ordered-set-lattice\ S\ n)$
 $\langle proof \rangle$

lemma *map2-max-mono*:

fixes $xs :: ('a :: linorder)\ list$
assumes $length\ xs = length\ ys$
assumes *sorted-wrt* $(<)\ xs$ *sorted-wrt* $(<)\ ys$
shows *sorted-wrt* $(<)\ (map2\ max\ xs\ ys)$
 $\langle proof \rangle$

lemma *map2-min-mono*:

fixes $xs :: ('a :: linorder)\ list$
assumes $length\ xs = length\ ys$
assumes *sorted-wrt* $(<)\ xs$ *sorted-wrt* $(<)\ ys$
shows *sorted-wrt* $(<)\ (map2\ min\ xs\ ys)$
 $\langle proof \rangle$

lemma *ordered-set-lattice-carrier-finite-ne*:

assumes $finite\ S\ n \leq card\ S$
shows $carrier\ (ordered-set-lattice\ S\ n) \neq \{\}$ *finite* $(carrier\ (ordered-set-lattice\ S\ n))$
 $\langle proof \rangle$

lemma *ordered-set-lattice-lattice*:

fixes $S :: ('a :: linorder)\ set$
assumes $finite\ S\ n \leq card\ S$
shows *finite-ne-distrib-lattice* $(ordered-set-lattice\ S\ n)$
 $\langle proof \rangle$

lemma *insort-eq*:

fixes $xs :: ('a :: linorder)\ list$

assumes *sorted xs*
shows $\exists ys zs. \text{insort } e \text{ } xs = ys @ e \# zs \wedge ys @ zs = xs \wedge \text{set } ys \subseteq \{..<e\} \wedge \text{set } zs \subseteq \{e..\}$
 $\langle \text{proof} \rangle$

lemma *list-all2-insort*:
fixes $xs \ ys :: ('a :: \text{linorder}) \text{ list}$
assumes $\text{length } xs = \text{length } ys \ \text{sorted } xs \ \text{sorted } ys$
shows $\text{list-all2 } (\leq) \ xs \ ys \longleftrightarrow \text{list-all2 } (\leq) \ (\text{insort } e \ xs) \ (\text{insort } e \ ys)$
 $\langle \text{proof} \rangle$

lemma *le-ordered-set-lattice-diff*:
fixes $x \ y :: ('a :: \text{linorder}) \text{ set}$
assumes $\text{finite } x \ \text{finite } y \ \text{card } x = \text{card } y$
shows $\text{le-ordered-set-lattice } x \ y \longleftrightarrow \text{le-ordered-set-lattice } (x - y) \ (y - x)$
 $\langle \text{proof} \rangle$

lemma *ordered-set-lattice-carrier*:
assumes $T \in \text{carrier } (\text{ordered-set-lattice } S \ n)$
shows $\text{finite } T \ \text{card } T = n \ T \subseteq S$
 $\langle \text{proof} \rangle$

lemma *ordered-set-lattice-dual*:
assumes $\text{finite } S \ n \leq \text{card } S$
defines $L \equiv \text{ordered-set-lattice } S \ n$
defines $M \equiv \text{ordered-set-lattice } S \ (\text{card } S - n)$
shows
 $\bigwedge x. x \in \text{carrier } L \implies (S - x) \in \text{carrier } M$
 $\bigwedge x. x \in \text{carrier } M \implies (S - x) \in \text{carrier } L$
 $\bigwedge x \ y. x \in \text{carrier } L \wedge y \in \text{carrier } L \implies x \sqsubseteq_L y \longleftrightarrow (S - y) \sqsubseteq_M (S - x)$
 $\langle \text{proof} \rangle$

lemma *bij-betw-ord-set-lattice-pairs*:
assumes $\text{finite } S \ n \leq \text{card } S$
defines $L \equiv \text{ordered-set-lattice } S \ n$
assumes $x \in \text{carrier } L \ y \in \text{carrier } L \ x \sqsubseteq_L y$
shows $\exists \varphi. \text{bij-betw } \varphi \ x \ y \wedge \text{strict-mono-on } x \ \varphi \wedge (\forall e. \varphi \ e \geq e)$
 $\langle \text{proof} \rangle$

definition $\text{bij-pmf } I \ F = \text{pmf-of-set } \{f. \text{bij-betw } f \ I \ F \wedge f \in \text{extensional } I\}$

lemma *card-bijections'*:
assumes $\text{finite } A \ \text{finite } B \ \text{card } A = \text{card } B$
shows $\text{card } \{f. \text{bij-betw } f \ A \ B \wedge f \in \text{extensional } A\} = \text{fact } (\text{card } A) \ (\text{is } ?L = ?R)$
 $\langle \text{proof} \rangle$

lemma *bij-betw-non-empty-finite*:
assumes $\text{finite } I \ \text{finite } F \ \text{card } I = \text{card } F$

shows
 $\text{finite } \{f. \text{bij-betw } f \ I \ F \wedge f \in \text{extensional } I\}$ (is ?T1)
 $\{f. \text{bij-betw } f \ I \ F \wedge f \in \text{extensional } I\} \neq \{\}$ (is ?T2)
 <proof>

lemma *bij-pmf*:
assumes $\text{finite } I \ \text{finite } F \ \text{card } I = \text{card } F$
shows
 $\text{set-pmf } (\text{bij-pmf } I \ F) = \{f. \text{bij-betw } f \ I \ F \wedge f \in \text{extensional } I\}$
 $\text{finite } (\text{set-pmf } (\text{bij-pmf } I \ F))$
 <proof>

lemma *expectation-ge-eval-at-point*:
assumes $\bigwedge y. y \in \text{set-pmf } p \implies f \ y \geq (0::\text{real})$
assumes *integrable* $p \ f$
shows $\text{pmf } p \ x * f \ x \leq (\int x. f \ x \ \partial p)$ (is ?L ≤ ?R)
 <proof>

lemma *split-bij-pmf*:
assumes $\text{finite } I \ \text{finite } F \ \text{card } I = \text{card } F \ J \subseteq I$
shows $\text{bij-pmf } I \ F =$
 do {
 $S \leftarrow \text{pmf-of-set } \{S. \text{card } S = \text{card } J \wedge S \subseteq F\};$
 $\varphi \leftarrow \text{bij-pmf } J \ S;$
 $\psi \leftarrow \text{bij-pmf } (I - J) \ (F - S);$
 $\text{return-pmf } (\text{merge } J \ (I - J) \ (\varphi, \psi))$
 } (is ?L = ?R)
 <proof>

lemma *map-bij-pmf*:
assumes $\text{finite } I \ \text{finite } F \ \text{card } I = \text{card } F \ \text{inj-on } \varphi \ F$
shows $\text{map-pmf } (\lambda f. (\lambda x \in I. \varphi(f \ x))) \ (\text{bij-pmf } I \ F) = \text{bij-pmf } I \ (\varphi \ ' \ F)$
 <proof>

lemma *pmf-of-multiset-eq-pmf-of-setI*:
assumes $c > 0 \ x \neq \{\#\}$
assumes $\bigwedge i. i \in y \implies \text{count } x \ i = c$
assumes $\bigwedge i. i \in \# \ x \implies i \in y$
shows $\text{pmf-of-multiset } x = \text{pmf-of-set } y$
 <proof>

lemma *card-multi-bij*:
assumes *finite* J
assumes $I = \bigcup (A \ ' \ J) \ \text{disjoint-family-on } A \ J$
assumes $\bigwedge j. j \in J \implies \text{finite } (A \ j) \wedge \text{finite } (B \ j) \wedge \text{card } (A \ j) = \text{card } (B \ j)$
shows $\text{card } \{f. (\forall j \in J. \text{bij-betw } f \ (A \ j) \ (B \ j)) \wedge f \in \text{extensional } I\} = \prod_{i \in J. \text{fact } (\text{card } (A \ i))}$
 (is card ?L = ?R)
 <proof>

lemma *map-bij-pmf-non-inj*:

fixes $I :: 'a \text{ set}$

fixes $F :: 'b \text{ set}$

fixes $\varphi :: 'b \Rightarrow 'c$

assumes $\text{finite } I \text{ finite } F \text{ card } I = \text{card } F$

defines $q \equiv \{f. f \in \text{extensional } I \wedge \{\#\# f x. x \in \#\ \text{mset-set } I\#\} = \{\#\# \varphi x. x \in \#\ \text{mset-set } F\#\}\}$

shows $\text{map-pmf } (\lambda f. (\lambda x \in I. \varphi(f x))) (\text{bij-pmf } I F) = \text{pmf-of-set } q \text{ (is ?L = -)}$
<proof>

lemmas $\text{fkg-inequality-pmf-internalized} = \text{fkg-inequality-pmf}[\text{unoverload-type } 'a]$

lemma *permutation-distributions-are-neg-associated*:

fixes $F :: ('a :: \text{linorder-topology}) \text{ set}$

fixes $I :: 'b \text{ set}$

assumes $\text{finite } F \text{ finite } I \text{ card } I = \text{card } F$

shows $\text{measure-pmf.neg-assoc } (\text{bij-pmf } I F) (\lambda i \omega. \omega i) I$
<proof>

lemma *multiset-permutation-distributions-are-neg-associated*:

fixes $F :: ('a :: \text{linorder-topology}) \text{ multiset}$

fixes $I :: 'b \text{ set}$

assumes $\text{finite } I \text{ card } I = \text{size } F$

defines $p \equiv \text{pmf-of-set } \{\varphi. \varphi \in \text{extensional } I \wedge \text{image-mset } \varphi (\text{mset-set } I) = F\}$

shows $\text{measure-pmf.neg-assoc } p (\lambda i \omega. \omega i) I$
<proof>

lemma *n-subsets-prob*:

assumes $d \leq \text{card } S \text{ finite } S \ s \in S$

shows

$\text{measure-pmf.prob } (\text{pmf-of-set } \{a. a \subseteq S \wedge \text{card } a = d\}) \{\omega. s \notin \omega\} = (1 - \text{real } d / \text{card } S)$

$\text{measure-pmf.prob } (\text{pmf-of-set } \{a. a \subseteq S \wedge \text{card } a = d\}) \{\omega. s \in \omega\} = \text{real } d / \text{card } S$

<proof>

lemma *n-subsets-distribution-neg-assoc*:

assumes $\text{finite } S \ k \leq \text{card } S$

defines $p \equiv \text{pmf-of-set } \{T. T \subseteq S \wedge \text{card } T = k\}$

shows $\text{measure-pmf.neg-assoc } p (\in) S$
<proof>

end

7 Application: Bloom Filters

The false positive probability of Bloom Filters is a case where negative association is really useful. Traditionally it is derived only approximately. Bloom [4] first derives the expected number of bits set to true given the number of elements inserted, then the false positive probability is computed, pretending that the expected number of bits is the actual number of bits.

Both Blooms original derivation and Mitzenmacher and Upfal [15] use this method.

A more correct approach would be to derive a tail bound for the number of set bits and derive a false-positive probability based on that, which unfortunately leads to a complex formula.

An exact result has later been derived using combinatorial methods by Gopinathan and Sergey [10]. However their formula is less useful, as it consists of a sum with Stirling numbers and binomial coefficients.

It is however easy to see that the original bound derived by Bloom is a correct upper bound for the false positive probability using negative association. (This is pointed out by Bao et al. [?].)

In this section, we derive the same bound using this library as an example for the applicability of this library.

theory *Negative-Association-Bloom-Filters*

imports *Negative-Association-Permutation-Distributions*

begin

fun *bloom-filter-pmf* **where**

```

bloom-filter-pmf 0 d N = return-pmf {} |
bloom-filter-pmf (Suc n) d N = do {
  h ← bloom-filter-pmf n d N;
  a ← pmf-of-set {a. a ⊆ {..N::nat} ∧ card a = d};
  return-pmf (a ∪ h)
}

```

lemma *bloom-filter-neg-assoc*:

assumes $d \leq N$

shows *measure-pmf.neg-assoc* (*bloom-filter-pmf* n d N) ($\lambda i \omega. i \in \omega$) {..*N*}
<proof>

lemma *bloom-filter-cell-prob*:

assumes $d \leq N$ $i < N$

shows *measure* (*bloom-filter-pmf* n d N) { $\omega. i \in \omega$ } = $1 - (1 - \text{real } d / \text{real } N)^n$
<proof>

lemma *bloom-filter-false-positive-prob*:

assumes $d \leq N$ $T \subseteq \{..*N*\}$ *card* T = d

shows *measure* (*bloom-filter-pmf* n d N) { $\omega. T \subseteq \omega$ } ≤ $(1 - (1 - \text{real } d / \text{real } N)^n)$

$N)^{\wedge n})^{\wedge d}$
(is $?L \leq ?R$)
<proof>

end

References

- [1] R. Ahlswede and D. E. Daykin. An inequality for the weights of two families of sets, their unions and intersections. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 43:183–185, 1978.
- [2] N. Alon and J. H. Spencer. *The Probabilistic Method, Second Edition*. John Wiley & Sons, Ltd, 2nd edition, 2000.
- [3] G. Birkhoff. *Lattice Theory*. AMS, 3rd edition, 1967.
- [4] B. H. Bloom. Space/time trade-offs in hash coding with allowable errors. *Commun. ACM*, 13(7):422–426, July 1970.
- [5] M. Doty. Birkhoff’s representation theorem for finite distributive lattices. *Archive of Formal Proofs*, December 2022. https://isa-afp.org/entries/Birkhoff_Finite_Distributive_Lattices.html, Formal proof development.
- [6] D. Dubhashi, J. Jonasson, and D. Ranjan. Positive influence and negative dependence. *Combinatorics, Probability and Computing*, 16(1):29–41, 2007.
- [7] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Structures & Algorithms*, 13(2):99–124, 1998.
- [8] D. P. Dubhashi, V. Priebe, and D. Ranjan. Negative dependence through the fkg inequality. *BRICS Report Series*, 3, 1996.
- [9] C. Fortuin, P. Kastelyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.*, 22:89–103, jun 1971.
- [10] K. Gopinathan and I. Sergey. Certifying certainty and uncertainty in approximate membership query structures. In S. K. Lahiri and C. Wang, editors, *Computer Aided Verification*, pages 279–303, Cham, 2020. Springer International Publishing.
- [11] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

- [12] R. Impagliazzo and V. Kabanets. Constructive proofs of concentration bounds. In M. Serna, R. Shaltiel, K. Jansen, and J. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 617–631, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [13] K. Joag-Dev and F. Proschan. Negative association of random variables with applications. *Annals of Statistics*, 11:286–295, 1983.
- [14] S. Lisawadi and T.-C. Hu. On the negative association property for the dependent bootstrap random variables. *Lobachevskii Journal of Mathematics*, 32:32–38, 2011.
- [15] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, USA, 2nd edition, 2017.
- [16] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [17] R. Pemantle. Towards a theory of negative dependence. *Journal of Mathematical Physics*, 41(3):1371–1390, 03 2000.