Verification of Discrete Upper Confidence Bound algorithm in Isabelle/HOL

Arjan Faber

December 26, 2025

Abstract

This project formally verifies the Upper Confidence Bound (UCB) algorithm in Isabelle/Higher-order Logic (HOL), focusing on its probabilistic guarantees and regret bounds. The work extends Isabelle/HOLs probabilistic framework and explores verification of discrete-time bandit models following [1]. This research advances the formal verification of probabilistic algorithms in reinforcement learning.

theory *MSc-Project-Discrete-Prop15-1*
  imports
    HOL−Probability.Probability

begin

locale *bandit-problem* =
  fixes A :: $'a$ *set*
    and μ :: $'a \Rightarrow$ *real*
    and *a-star* :: $'a$
  assumes *finite-arms*: *finite* A
    and *a-star-in-A*: *a-star* $\in A$
    and *optimal-arm*: $\forall\, a \in A.\ \mu$ *a-star* $\geq \mu\ a$
begin

definition Δ :: $'a \Rightarrow$ *real* **where**
  $\Delta\ a = \mu$ *a-star* $- \mu\ a$
end

locale *bandit-policy* = *bandit-problem* + *prob-space* +
  fixes Ω :: $'b$ *set*
    and $\mathcal{F}$ :: $'b$ *set set*
    and π :: *nat* $\Rightarrow$ $'b \Rightarrow$ $'a$
    and *N-n* :: *nat* $\Rightarrow$ $'a \Rightarrow$ $'b \Rightarrow$ *nat*
  assumes *measurable-policy*: $\forall\, t.\ \pi\ t \in$ *measurable* M (*count-space* A)
    and *N-n-def*: $\forall\, n\ a\ \omega.$ *N-n* $n\ a\ \omega =$ *card* $\{t \in \{0..<n\}.\ \pi\ (t{+}1)\ \omega = a\}$
    and *count-assm-pointwise*: $\forall\, n\ \omega.\ (\sum a \in A.$ *real* (*N-n* $n\ a\ \omega$)) = *real* n

begin

definition *R-n* :: *nat* ⇒ *'b* ⇒ *real* **where**
  R-n n ω = *n* ∗ *μ a-star* − (∑ *a* ∈ *A*. *μ a* ∗ *real* (*N-n n a ω*))

lemma *regret-decomposition-pointwise*:
  fixes *n* :: *nat* **and** *ω* :: *'b*
  assumes *n-count-assm-pointwise*: (∑ *a*∈*A*. *real* (*N-n n a ω*)) = *real n*
  shows *R-n n ω* = (∑ *a* ∈ *A*. Δ *a* ∗ *real* (*N-n n a ω*))
proof −
  have *sum-Nn*: (∑ *a* ∈ *A*. *real* (*N-n n a ω*)) = *real n*
    using *n-count-assm-pointwise* **by** *simp*

  have *sum-const*: (∑ *a* ∈ *A*. *μ a-star* ∗ *real* (*N-n n a ω*)) = *μ a-star* ∗ (∑ *a* ∈ *A*. *real* (*N-n n a ω*))
    by (*simp add*: *Cartesian-Space.vector-space-over-itself.scale-sum-right*)

  have *eq1*: *R-n n ω* = *real n* ∗ *μ a-star* − (∑ *a* ∈ *A*. *μ a* ∗ *real* (*N-n n a ω*))
    by (*simp add*: *R-n-def*)

  also have *eq2*: ... = (∑ *a* ∈ *A*. *μ a-star* ∗ *real* (*N-n n a ω*)) − (∑ *a* ∈ *A*. *μ a* ∗ *real* (*N-n n a ω*))
    using *sum-const sum-Nn*
    by (*subst sum-Nn*[*symmetric*], *subst sum-const*) *simp*

  also have *eq3*: ... = (∑ *a* ∈ *A*. (*μ a-star* ∗ *real* (*N-n n a ω*) − *μ a* ∗ *real* (*N-n n a ω*)))
    by (*rule sum-subtractf*[*symmetric*])

  also have *eq4*: ... = (∑ *a* ∈ *A*. (*μ a-star* − *μ a*) ∗ *real* (*N-n n a ω*))
    by (*simp add*: *algebra-simps*)

  also have ... = (∑ *a* ∈ *A*. Δ *a* ∗ *real* (*N-n n a ω*))
    by (*simp add*: Δ*-def*)

  finally show *?thesis* **.**
qed

lemma *integrable-const-fun*:
  assumes *finite-measure M*
  shows *integrable M* (*λx*. *c*)
  using *assms* **by** (*simp add*: *Bochner-Integration.finite-measure.integrable-const*)

lemma *expected-regret*:
  assumes *finite A*
  and ∀ *a* ∈ *A*. *integrable M* (*λω*. *real* (*N-n n a ω*))
  shows *expectation* (*λω*. *R-n n ω*) = (∑ *a*∈*A*. Δ *a* ∗ *expectation* (*λω*. *real* (*N-n n a ω*)))

```
proof -

  have integrable-sum: integrable M (λω. ∑ a ∈ A. μ a * real (N-n n a ω))
  proof -
    have ∀ a ∈ A. integrable M (λω. μ a * real (N-n n a ω))
      using assms integrable-mult-right by blast
    then show ?thesis
      by (simp add: integrable-sum)
  qed

  have pointwise: ∀ ω. R-n n ω = (∑ a∈A. Δ a * real (N-n n a ω))
    using regret-decomposition-pointwise count-assm-pointwise by blast

  hence rewrite: expectation (λω. R-n n ω) = expectation (λω. ∑ a∈A. Δ a * real
(N-n n a ω))
    by simp

   also have expectation (λω. R-n n ω) = (∑ a∈A. Δ a * expectation (λω. real
(N-n n a ω)))
  proof -

    have ∀ a ∈ A. integrable M (λω. Δ a * real (N-n n a ω))
    using assms integrable-mult-right by blast
    hence integrable M (λω. ∑ a ∈ A. Δ a * real (N-n n a ω))
     using integrable-sum by simp


     have res-1: expectation (λω. ∑ a∈A. Δ a * real (N-n n a ω)) = integral^L M
(λω. ∑ a∈A. Δ a * real (N-n n a ω))
       by simp

     also have res-2: ... = (∑ a∈A. integral^L M (λω. Δ a * real (N-n n a ω)))
       using ‹finite A› assms
       by (simp add: integral-sum integrable-sum)

     have res-3: expectation (λω. R-n n ω) = (∑ a∈A. Δ a * expectation (λω. real
(N-n n a ω)))
      using assms rewrite
      by simp

    then show ?thesis using assms res-3 by auto
  qed

  then show ?thesis .
qed



end
```

```
end
theory Discrete-UCB-Step1
  imports MSc-Project-Discrete-Prop15-1

begin

locale bandit-policy = bandit-problem +  prob-space  +
  fixes Ω :: 'b set
    and ℱ :: 'b set set
    and ω :: 'b
    and π :: nat ⇒ 'b ⇒ 'a
    and N-n :: nat ⇒ 'a ⇒ 'b ⇒ nat
  assumes  measurable-policy: ∀ t. π t ∈ measurable M (count-space A)
    and N-n-def: ∀ n a ω. N-n n a ω = card {t ∈ {0..<n}. π (t+1) ω = a}
    and count-assm-pointwise: ∀ n ω. (∑ a ∈ A. real (N-n n a ω)) = real n
begin

lemma union-eq:
  fixes a :: 'a and n k :: nat
  assumes k ≤ n
  shows {t. t < n ∧ π (t+1) ω = a} = {t. t < k ∧ π (t+1) ω = a} ∪ {t. k ≤ t
∧ t < n ∧ π (t+1) ω = a}
proof
  show {t. t < n ∧ π (t+1) ω = a} ⊆ {t. t < k ∧ π (t+1) ω = a} ∪ {t. k ≤ t ∧
t < n ∧ π (t+1) ω = a}
  proof
    fix x assume x ∈ {t. t < n ∧ π (t+1) ω = a}
    then have x < n and π (x+1) ω = a by auto
    then have x < k ∨ k ≤ x by auto
    thus x ∈ {t. t < k ∧ π (t+1) ω = a} ∪ {t. k ≤ t ∧ t < n ∧ π (t+1) ω = a}
      using ‹x < n› ‹π (x+1) ω = a› by auto
  qed
next
  show {t. t < k ∧ π (t+1) ω = a} ∪ {t. k ≤ t ∧ t < n ∧ π (t+1) ω = a} ⊆ {t.
t < n ∧ π (t+1) ω = a}
  proof
    fix x assume x ∈ {t. t < k ∧ π (t+1) ω = a} ∪ {t. k ≤ t ∧ t < n ∧ π (t+1)
ω = a}
    then have x < k ∧ π (x+1) ω = a ∨ (k ≤ x ∧ x < n ∧ π (x+1) ω = a) by
auto
    hence x < n ∧ π (x+1) ω = a using assms by auto
    thus x ∈ {t. t < n ∧ π (t+1) ω = a} by auto
  qed
qed

lemma cardinality-indic-eq:
  fixes I :: nat ⇒ bool
  assumes finite {t. k ≤ t ∧ t < n}
```

shows *card* $\{t.\ k \le t \wedge t < n \wedge \pi\ (t{+}1)\ \omega = a \wedge I\ t\} = (\sum\ t = k..{<}n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\ \mathit{then}\ 1\ \mathit{else}\ 0)$
proof −
have *fin*: *finite* $\{t.\ k \le t \wedge t < n\}$ **using** *assms* **by** *simp*
have *sum-eq*:
sum (*indicator* $\{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$) $\{t.\ k \le t \wedge t < n\}$ = *card* ($\{t.\ k \le t \wedge t < n\} \cap \{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$)
by (*rule sum-indicator-eq-card*[*OF fin*])
have *sets-eq*:
$\{t.\ k \le t \wedge t < n\} \cap \{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\} = \{t.\ k \le t \wedge t < n \wedge \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$
by *auto*
hence *card-eq*:
card ($\{t.\ k \le t \wedge t < n\} \cap \{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$) = *card* $\{t.\ k \le t \wedge t < n \wedge \pi(t{+}1)\ \omega = a \wedge I\ t\}$
by *simp*
have *card-sum-eq*:
sum (*indicator* $\{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$) $\{t.\ k \le t \wedge t < n\}$ = *card* $\{t.\ k \le t \wedge t < n \wedge \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$
using *sum-eq sets-eq* **by** *simp*
have *ind-eq*:
sum (*indicator* $\{t.\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\}$) $\{t.\ k \le t \wedge t < n\}$ = *sum* ($\lambda t.$ *of-bool* ($\pi\ (t{+}1)\ \omega = a \wedge I\ t$)) $\{t.\ k \le t \wedge t < n\}$
by (*simp add*: *indicator-def*)
have *bool-if-eq*:
sum ($\lambda t.$ *of-bool* ($\pi\ (t{+}1)\ \omega = a \wedge I\ t$)) $\{t.\ k \le t \wedge t < n\}$ = *sum* ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) $\{t.\ k \le t \wedge t < n\}$
by (*simp add*: *of-bool-def*)
have *set-eq*: $\{t.\ k \le t \wedge t < n\} = \{k..\} \cap \{..{<}n\}$ **by** *auto*
have *sum-set-eq*:
sum ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) $\{t.\ k \le t \wedge t < n\}$ = *sum* ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) ($\{k..\} \cap \{..{<}n\}$)
by (*simp add*: *set-eq*)
have *atLeastLessThan-eq*:
sum ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) ($\{k..\} \cap \{..{<}n\}$) = *sum* ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) (*atLeastLessThan k n*)
by (*simp add*: *atLeastLessThan-def*)
have *final-sum-eq*:
sum ($\lambda t.$ *if* $\pi\ (t{+}1)\ \omega = a \wedge I\ t$ *then 1 else 0*) (*atLeastLessThan k n*) = $(\sum\ t = k..{<}n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge I\ t\ \mathit{then}\ 1\ \mathit{else}\ 0)$
by *simp*
show *?thesis*
apply (*subst card-sum-eq*[*symmetric*])
apply (*subst ind-eq*)
apply (*subst bool-if-eq*)
apply (*subst sum-set-eq*)
apply (*subst atLeastLessThan-eq*)
apply (*simp add*: *final-sum-eq*)
done

qed

lemma *ge-rewrite*: $(x{::}real) \geq y \Longrightarrow y \leq x$ **by** *simp*

lemma *Nn-expression*:
  fixes $a :: 'a$ **and** $s :: nat \Rightarrow real$
    and $k :: nat$ **and** $n :: nat$
  assumes $a \in A$
    and $k \leq n$
    and $0 < n$
    and $\forall\ t \in \{0..n\}.\ 0 < s\ t$
    and $\forall t < n - 1.\ \ s\ t \leq s\ (t + 1)$
    and *init-play-once*: $\forall \omega.\ a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$
    and *finite-played-sets*:
    $finite\ \{t.\ t < n \wedge \pi\ (t{+}1)\ \omega = a\}$
    $finite\ \{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\}$
    $finite\ \{t.\ k \leq t \wedge t < n \wedge \pi\ (t{+}1)\ \omega = a\}$
  shows
    $(N\text{-}n\ n\ a\ \omega) = 1 + (\sum\ t = k..{<}n.\ if\ \pi\ (t{+}1)\ \omega = a \wedge real\ (N\text{-}n\ t\ a\ \omega) < s\ t$ $then\ 1\ else\ 0) +$
                    $(\sum\ t = k..{<}n.\ if\ \pi\ (t{+}1)\ \omega = a \wedge real\ (N\text{-}n\ t\ a\ \omega) \geq s\ t\ then\ 1$ $else\ 0)$
proof –
  have *Nn-def*: $(N\text{-}n\ n\ a\ \omega) = card\ \{t.\ t < n \wedge \pi\ (t{+}1)\ \omega = a\}$
    by (*simp add*: *N-n-def*)

  have *init-count*: $1 \leq card\ \{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\}$
  proof –
    have *eq-card*: $N\text{-}n\ k\ a\ \omega = card\ \{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\}$
      by (*simp add*: *N-n-def*)
    moreover from *init-play-once* **and** ‹$a \in A$› **have** $1 \leq N\text{-}n\ k\ a\ \omega$
      by *simp*
    ultimately show *?thesis* **by** *simp*
  qed

  have *disj*: $\{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\} \cap \{t.\ k \leq t \wedge t < n \wedge \pi\ (t{+}1)\ \omega = a\}$ $= \{\}$
    by *auto*

  have *union-eq-set*:
    $\{t.\ t < n \wedge \pi\ (t{+}1)\ \omega = a\} = \{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\} \cup \{t.\ k \leq t \wedge t <$ $n \wedge \pi\ (t{+}1)\ \omega = a\}$
    using *union-eq assms* **by** *blast*

  have *finite1*: $finite\ \{t.\ t < k \wedge \pi\ (t{+}1)\ \omega = a\}$ **and** *finite2*: $finite\ \{t.\ k \leq t \wedge t$ $< n \wedge \pi\ (t{+}1)\ \omega = a\}$
    using *finite-played-sets* **by** *auto*

  have *card-eq*:

card $\{t.\ t < n \wedge \pi\ (t+1)\ \omega = a\} = $ *card* $\{t.\ t < k \wedge \pi\ (t+1)\ \omega = a\} +$ *card* $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\}$
using *card-Un-disjoint*[*OF finite1 finite2 disj*] *union-eq-set* **by** *simp*

have *eq-card-k*: *card* $\{t.\ t < k \wedge \pi\ (t+1)\ \omega = a\} = $ *N-n k a* ω
by (*simp add*: *N-n-def*)

have *card-k-eq-1*: *card* $\{t.\ t < k \wedge \pi\ (t+1)\ \omega = a\} = 1$
using *init-play-once assms eq-card-k* **by** *simp*

have *card-eq-1*:
card $\{t.\ t < n \wedge \pi\ (t+1)\ \omega = a\} = 1 +$ *card* $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\}$
using *card-eq eq-card-k card-k-eq-1* **by** *simp*

have *finite-lt*: *finite* $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t\}$
using *finite2 finite-subset* **by** *simp*

have *finite-ge*: *finite* $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $\geq s\ t\}$
using *finite2 finite-subset* **by** *simp*

have *card-eq-partition*:
card $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\} =$
card $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t\} +$
card $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $\geq s\ t\}$
proof −
have *union-eq-partition*:
$\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\} =$
$\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t\} \cup$
$\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $\geq s\ t\}$ **by** *auto*

have *disj-partition*:
$\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t\} \cap$
$\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $\geq s\ t\} = \{\}$
by *auto*

show *?thesis*
using *card-Un-disjoint*[*OF finite-lt finite-ge disj-partition*] *union-eq-partition* **by** *simp*
qed

have *card-eq-sum-lt*:
card $\{t.\ k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t\} =$
($\sum$ *t = k..<n. if* $\pi\ (t+1)\ \omega = a \wedge$ *real* (*N-n t a* ω) $< s\ t$ *then 1 else 0*)
using *cardinality-indic-eq* **by** *simp*

have *card-eq-sum-ge*:

```
    card {t. k ≤ t ∧ t < n ∧ π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t} =
     (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t then 1 else 0)
    using cardinality-indic-eq by simp

  have (N-n n a ω) = card {t. t < n ∧ π (t+1) ω = a}
    using N-n-def by simp

  also have ... = 1 +
    (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) < s t then 1 else 0) +
    (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t then 1 else 0)
    using card-eq-partition card-eq-1 card-eq-sum-lt card-eq-sum-ge by simp

  also have ... = 1 +
    (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) < s t then 1 else 0) +
    (∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0)
    using ge-rewrite by simp

  have rewrite-eq: (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t then
1 else 0) =
       (∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0)
    using sum.cong by simp

  ultimately have final-eq: (N-n n a ω) = 1 + (∑ t = k..<n. if π (t+1) ω = a
∧ real (N-n t a ω) < s t then 1 else 0) +
                      (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t then 1
else 0)
    by simp

  show (N-n n a ω) = 1 + (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω)
< s t then 1 else 0) +
                      (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) ≥ s t then 1
else 0)
    apply (subst rewrite-eq)
    apply (subst final-eq)
    apply simp
    done

qed

lemma upper-bound-expression-contradiction:
  fixes a :: 'a and s :: nat ⇒ real
    and k :: nat and n :: nat
    and s-n-nat :: nat
  assumes a ∈ A
    and k ≤ n
    and 0 < n
    and non-neg-s: ∀ t ∈ {0..n}. 0 < s t
    and base-le: s 0 ≤ s 1
    and non-dec: ∀t < n − 1. s t ≤ s (t + 1)
```

and *s-mono*: $\bigwedge t.\ k \leq t \land t \leq n \Longrightarrow s\ t \leq s\ n$
and *init-play-once*: $\forall \omega.\ a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$
and *finite-played-sets*:
$finite\ \{t.\ t < n \land \pi\ (t+1)\ \omega = a\}$
$finite\ \{t.\ t < k \land \pi\ (t+1)\ \omega = a\}$
$finite\ \{t.\ k \leq t \land t < n \land \pi\ (t+1)\ \omega = a\}$
and *xs-sorted-def*: $xs = sorted\text{-}list\text{-}of\text{-}set\ \{t \in \{k..<n\}.\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\}$
and *s-nat-def*: $s\text{-}n\text{-}nat = nat\ (\lfloor s\ n \rfloor)$
and *len-bound-def*: $s\text{-}n\text{-}nat < length\ xs$
and *distinct-xs*: $distinct\ xs$
and *gt-ineq*: $length\ xs + 1 > \lfloor s\ n \rfloor$
and *N-n-increasing-with-plays*:
$\forall\ t\ t'.\ k \leq t \land t < t' \land \pi\ (t+1)\ \omega = a \land \pi\ (t'+1)\ \omega = a \longrightarrow N\text{-}n\ t'\ a\ \omega \geq N\text{-}n\ t\ a\ \omega + 1$
and *neg*: $1 + real\ (\sum\ t=k..<n.\ if\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\ then\ 1\ else\ 0) > s\ n$
and $t\text{-}hat \in set\ xs$
and $t\text{-}hat = xs\ !\ s\text{-}n\text{-}nat$
and $real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq real\ s\text{-}n\text{-}nat + 1$

shows $(real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq \lfloor s\ n \rfloor + 1) \land (\pi\ (t\text{-}hat+1)\ \omega = a \land (real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) < s\ t\text{-}hat))$

proof –
have *sn-nat-def*:$s\text{-}n\text{-}nat = nat\ (\lfloor s\ n \rfloor)$ **by** *fact*

let $?I = \{t \in \{k..<n\}.\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\}$
have *sum-eq-card*: $real\ (\sum\ t=k..<n.\ if\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\ then\ 1\ else\ 0) = real\ (card\ ?I)$
using *cardinality-indic-eq* **by** *simp*

have *finI*: $finite\ \{t \in \{k..<n\}.\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\}$
proof –
have *subset*: $\{t \in \{k..<n\}.\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\} \subseteq \{t.\ k \leq t \land t < n \land \pi\ (t+1)\ \omega = a\}$
by *auto*
moreover have $finite\ \{t.\ k \leq t \land t < n \land \pi\ (t+1)\ \omega = a\}$
using *finite-played-sets(3)* **.**
ultimately show *?thesis*
using *finite-subset* **by** *blast*
qed

have *xs-props*: $sorted\ xs\ distinct\ xs\ set\ xs = \{t \in \{k..<n\}.\ \pi\ (t+1)\ \omega = a \land real\ (N\text{-}n\ t\ a\ \omega) < s\ t\}$
using *List.linorder-class.set-sorted-list-of-set*

using *finite-played-sets*
unfolding *xs-sorted-def*
by *auto*

have *sum-eq-length*:
$(\sum t = k..<n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t\ \mathit{then}\ 1\ \mathit{else}\ 0) = \mathit{length}\ xs$
proof –
let $?P = \lambda t.\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t$
let $?I = \{t \in \{k..<n\}.\ ?P\ t\}$

have *sum-card*: $(\sum t = k..<n.\ \mathit{if}\ ?P\ t\ \mathit{then}\ 1\ \mathit{else}\ 0) = \mathit{card}\ ?I$
using *cardinality-indic-eq* **by** *simp*

also have ... = $\mathit{card}\ (\mathit{set}\ xs)$
using *xs-props*(*3*) **by** *simp*

also have ... = $\mathit{length}\ xs$
using *xs-props*(*2*) *distinct-card* **by** *auto*

finally show *?thesis* **.**
qed

have *gt-ineq*: $1 + \mathit{real}(\mathit{length}\ xs) > s\ n$
proof –

have *eq1*:$1 + (\sum t = k..<n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t\ \mathit{then}\ 1\ \mathit{else}\ 0) = 1 + \mathit{length}\ xs$
using *sum-eq-length* **by** *simp*

then have *eq2*: $\mathit{real}\ (1 + (\sum t = k..<n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t\ \mathit{then}\ 1\ \mathit{else}\ 0)) = 1 + \mathit{real}\ (\ \mathit{length}\ xs)$
by *simp*

from *neg* **have** $\neg\ (1 + \mathit{real}\ (\sum t{=}k..<n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t\ \mathit{then}\ 1\ \mathit{else}\ 0) \le s\ n)$
by *simp*

hence *gt*: $1 + \mathit{real}\ (\sum t{=}k..<n.\ \mathit{if}\ \pi\ (t{+}1)\ \omega = a \wedge \mathit{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t\ \mathit{then}\ 1\ \mathit{else}\ 0) > s\ n$
by *simp*

then show *?thesis*
using *eq1 eq2 neg gt* **by** *simp*
qed

then have *len-bound-assm*: *s-n-nat* < *length xs*
  using *len-bound-def assms*
  by *auto*

then have *distinct-xs*: *distinct xs*
  using *assms* **by** *simp*

then have *gt-ineq*: *length xs* + *1* > ⌊*s n*⌋
  using *assms* **by** *simp*

have *part-1-glob*: π (*t-hat* +*1*) ω = *a* ∧ *real* (*N-n t-hat a* ω) < *s t-hat*
proof –
  have *floor-le-sn*: ⌊*s n*⌋ ≤ *s n*
    using *Archimedean-Field.of-int-floor-le* **by** *simp*

  from *gt-ineq* **have** *len-bound*: *length xs* + *1* > ⌊*s n*⌋
    using *len-bound-assm* **by** *simp*

  have *nat-expression-sn*: *length xs* ≥ *nat* (⌊*s n*⌋)
  proof –
    from *len-bound* **have** *length xs* ≥ *nat* (⌊*s n*⌋)
      by (*simp add*: *nat-less-iff*)
    then show *?thesis* **.**
  qed

  have *sn-lt-len-plus-1*: ⌊*s n*⌋ < *1* + *real* (*length xs*)
    using *gt-ineq* **by** *simp*

  then have *final-eq-glob*:π (*t-hat*+*1*) ω = *a* ∧ *real* (*N-n t-hat a* ω) < *s t-hat*
  proof –

    have π*-eq*: π (*t-hat*+*1*) ω = *a*
      using *assms xs-sorted-def* **by** *simp*

    then have π (*t-hat* +*1*) ω = *a* **using** π*-eq* **by** *simp*

    then have *intermid-eq*: *k* ≤ *t-hat* ∧ *t-hat* < *n* ∧ π (*t-hat* +*1*) ω = *a* ∧ *real* (*N-n t-hat a* ω) < *s t-hat*
      using *xs-sorted-def* ‹*t-hat* ∈ *set xs*› **by** *auto*

    have *fin-eq*:π (*t-hat* + *1*) ω = *a* ∧ *real* (*N-n t-hat a* ω) < *s t-hat*
      using *intermid-eq*
      by *simp*
    then show *?thesis* **by** *simp*
  qed

  show *?thesis* **using** *final-eq-glob* **by** *simp*

qed

have *floor-le-sn*: $\lfloor s\ n \rfloor \leq s\ n$
  using *Archimedean-Field.of-int-floor-le* **by** *simp*

have *one-plus-length-gt-floor*: $1 + real\ (length\ xs) > s\ n$
  using ‹$1 + real\ (length\ xs) > s\ n$› **.**

have *floor-less-length-plus-one*: $real\ (nat\ (\lfloor s\ n \rfloor)) < 1 + real\ (length\ xs)$
  using *one-plus-length-gt-floor floor-le-sn*
  by *linarith*

from *gt-ineq* **have** *len-result*: $length\ xs\ > \lfloor s\ n \rfloor - 1$ **by** *simp*
then have $length\ xs \geq nat\ (\lfloor s\ n \rfloor)$ **by** (*simp add*: *nat-less-iff*)
hence *t-hat-in-set*:$t\text{-}hat \in set\ xs$
  using *assms distinct-xs len-bound-assm* **by** *auto*

have *t-hat-le-n*: $t\text{-}hat + 1 \leq n$
  using *xs-sorted-def t-hat-in-set* **by** *auto*

have *k-le-t-hat*: $k \leq t\text{-}hat$
  using *xs-sorted-def t-hat-in-set* **by** *auto*

have *s-t-hat-le-s-n*: $s\ t\text{-}hat \leq s\ n$
  using *s-mono*[*of t-hat*]
  using *xs-sorted-def t-hat-in-set* **by** *auto*

have $real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq real\ s\text{-}n\text{-}nat + 1$
  using *assms* **by** *simp*

have *final-result*: $real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq \lfloor s\ n \rfloor + 1$
  using ‹$real\ (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq real\ (s\text{-}n\text{-}nat) + 1$›
  by (*simp add*: *sn-nat-def*)

show *?thesis* **using** *final-result* **and** *part-1-glob* **by** *auto*

qed

```
lemma Nn-upper-bound:
  fixes a :: 'a and s :: nat ⇒ real
    and k :: nat and n :: nat
  assumes asm: real(1 + (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) <
s t then 1 else 0)) ≤ s n
    and a-in-A: a ∈ A
    and k-le-n: k ≤ n
    and n-pos: 0 < n
    and s-pos: ∀ t ∈ {0..n}. 0 < s t
    and s-nondec: ∀ t < n − 1. s t ≤ s (t + 1)
    and init-play-once: ∀ω. a ∈ A ⟶ N-n k a ω = 1
    and finite-played-sets-1: finite {t. t < n ∧ π (t+1) ω = a}
    and finite-played-sets-2: finite {t. t < k ∧ π (t+1) ω = a}
    and finite-played-sets-3: finite {t. k ≤ t ∧ t < n ∧ π (t+1) ω = a}
  shows real(N-n n a ω) ≤ s n + real((∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤
real (N-n t a ω) then 1 else 0))
proof −
  have expr-nat: N-n n a ω =
    1 + (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t a ω) < s t then 1 else 0) +
        (∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0)
    using Nn-expression[OF a-in-A k-le-n n-pos s-pos s-nondec init-play-once
        finite-played-sets-1 finite-played-sets-2 finite-played-sets-3]
    by simp

  from asm have bound:real(1 + (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n t
a ω) < s t then 1 else 0)) ≤ s n by simp

  have real (N-n n a ω) = real ( 1 + (∑ t = k..<n. if π (t+1) ω = a ∧ real (N-n
t a ω) < s t then 1 else 0) +
        (∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0))
    using expr-nat by simp

  also have ... ≤ s n + real ( (∑ t = k..<n. if π (t+1) ω = a ∧ s t ≤ real (N-n
t a ω) then 1 else 0))
    using assms
    by simp

  then show ?thesis
    using expr-nat asm
    by simp
qed


theorem ENn-upper-bound:
  assumes
    a-in-A: a ∈ A
    and k-le-n: k ≤ n
    and n-pos: 0 < n
    and s-pos: ∀ t ∈ {0..n}. 0 < s t
```

and *s-nondec*: $\forall\, t < n.\ s\ t \leq s\ (t + 1)$
and *init-play-once*: $\forall\, \omega.\ a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$
and *integrable-Nn*: *integrable* $M\ (\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega))$
and *integrable-rhs-sum*: *integrable* $M\ (\lambda\omega.\ s\ n + (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0))$
and *integrable-s*: *integrable* $M\ (\lambda\omega.\ s\ n)$
and *integrable-indicator-sum*: *integrable* $M\ (\lambda\omega.\ \sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
and *linearity*: $integral^L\ M\ (\lambda\omega.\ s\ n + (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$integral^L\ M\ (\lambda\omega.\ s\ n) + integral^L\ M\ (\lambda\omega.\ \sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
and *pointwise-bound*: $real\ (N\text{-}n\ n\ a\ \omega) \leq s\ n + (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
and *mono-intgrl*: $integral^L\ M\ (\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)) \leq integral^L\ M\ (\lambda\omega.\ s\ n + (\sum t = k..<n.$
$if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0))$
shows
$expectation\ (\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)) \leq$
$s\ n + expectation\ (\lambda\omega.\ (\sum t = k..<n.\ \ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0))$
proof −
let $?f = \lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)$
let $?g = \lambda\omega.\ s\ n + (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
let $?g1 = \lambda\omega.\ s\ n$
let $?g2 = \lambda\omega.\ (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
have *pointwise-le*: $?f\ \omega \leq\ ?g\ \omega$
using *pointwise-bound* **.**

have *intg-f*: *integrable* $M\ ?f$ **using** *integrable-Nn* **.**
have *intg-g*: *integrable* $M\ ?g$ **using** *integrable-rhs-sum* **.**
then have *eq-f*: $integral^L\ M\ ?f = expectation\ (?f)$
using *prob-space* **by** *simp*
then have *eq-g*: $integral^L\ M\ ?g = expectation\ (?g)$
by *simp*

then have *sub-eqg1*: $integral^L\ M\ ?g1 = expectation\ (\lambda\omega.\ s\ n)$
by *simp*
then have *sub-eqg2*: $integral^L\ M\ ?g2 = expectation(\lambda\omega.\ (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0))$
by *simp*

have $real\ (N\text{-}n\ n\ a\ \omega) \leq s\ n + (\sum t = k..<n.\ if\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)$
using *assms* **by** *simp*

have *const-measure*: *Sigma-Algebra.measure M* (*space M*) = *1*
using *prob-space Probability-Measure.prob-space.prob-space* **by** *blast*

have *exp-sn*: *expectation* ($\lambda\omega$. *s n*) = *s n*
proof −
have *expectation* ($\lambda\omega$. *s n*) = *integral*^{L} *M* ($\lambda\omega$. *s n*)
by *simp*
also have ... = *s n*
using *prob-space* **by** (*simp add*: *prob-space.prob-space*)
then show *?thesis* **.**
qed

have *rhs-expr*:
integral^L *M ?g* = *integral*^{L} *M* ($\lambda\omega$. *s n*) + *integral*^{L} *M* ($\lambda\omega$. $\sum$ *t* = *k*..<*n*. *if* π (*t+1*) ω = *a* $\wedge$ *s t* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*)
using *linearity* **by** *auto*

then have *rhs-expr*:
expectation (*?g*) = *expectation* ($\lambda\omega$. *s n*) + *expectation* ($\lambda\omega$. ($\sum$ *t* = *k*..<*n*. *if* π (*t+1*) ω = *a* $\wedge$ *s t* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*))
by *simp*

also have *final-eq*: *expectation* (*?g*) = *s n* + *expectation* ($\lambda\omega$. ($\sum$ *t* = *k*..<*n*. *if* π (*t+1*) ω = *a* $\wedge$ *s t* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*))
using *rhs-expr exp-sn*
by *simp*

have *glob-final-eq*:*expectation* (*?f*) $\leq$ *expectation* (*?g*)
using *assms intg-f intg-g pointwise-le mono-intgrl*
by *auto*

thus *?thesis*
using *glob-final-eq final-eq* **by** *simp*
qed

end
end
theory *Discrete-UCB-Step2*
imports *Discrete-UCB-Step1*

begin

locale *bandit-policy* = *bandit-problem* + *prob-space* +
fixes Ω :: *'b set*
and $\mathcal{F}$ :: *'b set set*
and ω :: *'b*
fixes π :: *nat* $\Rightarrow$ *'b* $\Rightarrow$ *'a*
and *N-n* :: *nat* $\Rightarrow$ *'a* $\Rightarrow$ *'b* $\Rightarrow$ *nat*

and Z :: *nat* $\Rightarrow$ *'a* $\Rightarrow$ *'b* $\Rightarrow$ *real*
and δ :: *real*
and q :: *real*
assumes *finite-A*: *finite A*
and *a-in-A*: $a \in A$
and *measurable-policy*: $\forall\, t.\ \pi\ t \in$ *measurable M* (*count-space A*)
and *N-n-def*: $\forall\, n\ a\ b.$ *N-n n a b* = *card* $\{t \in \{0..<n\}.\ \pi\ (t+1)\ b = a\}$
and δ-*pos*: $0 < \delta$
and δ-*less1*: $\delta < 1$
and *q-pos*: $q > 0$

begin

definition *sample-mean-Z* :: *nat* $\Rightarrow$ *'a* $\Rightarrow$ *'b* $\Rightarrow$ *real* **where**
sample-mean-Z n a b $\equiv$ $(1\ /\ \mathit{real}\ n) * (\sum\ i{<}n.\ Z\ i\ a\ b)$

definition *M-val* :: *nat* $\Rightarrow$ *'a* $\Rightarrow$ *'b* $\Rightarrow$ *real* **where**
M-val t a b $\equiv$ (*if N-n* $(t+1)$ *a b* $= 0$ *then* 0
else $(\sum\ s < t.$ *if* $\pi\ s\ b = a$ *then* $Z\ s\ a\ b$ *else* $0)$ / *real* (*N-n t a b*))

definition U :: *nat* $\Rightarrow$ *'a* $\Rightarrow$ *'b* $\Rightarrow$ *real* **where**
$U\ t\ a\ b \equiv$ *M-val t a b* + q * *sqrt* (*ln* $(1\ /\ \delta)$ / $(2 *$ *real* (*max* 1 (*N-n t a b*))))

definition *A-t-plus-1* :: *nat* $\Rightarrow$ *'b* $\Rightarrow$ *'a* **where**
A-t-plus-1 t b $\equiv$ (*SOME a.* $a \in A \wedge (\forall\, a'.\ a' \in A \longrightarrow U\ t\ a\ b \geq U\ t\ a'\ b)$)

lemma (**in** *finite-measure*) *finite-measure-mono*:
assumes $A \subseteq B$ $B \in$*sets M* **shows** *measure M A* $\leq$ *measure M B*
using *emeasure-mono*[*OF assms*] *emeasure-real*[*of A*] *emeasure-real*[*of B*] **by** (*auto simp*: *measure-def*)

theorem *union-bound*:
fixes $E\ F\ G$ :: *'b set*
assumes $E \subseteq F \cup G$
and $E \in$ *events* $F \in$ *events* $G \in$ *events*
shows *prob E* $\leq$ *prob F* + *prob G*
proof −
have $F \cup G \in$ *events*
using *assms*(3,4) *sets.Un* **by** *blast*
have *prob E* $\leq$ *prob* $(F \cup G)$
using *assms local.finite-measure-mono* **by** *auto*
also have *prob* $(F \cup G) \leq$ *prob F* + *prob G*
using *assms measure-Un-le* **by** *blast*
finally show *?thesis* **.**
qed

```
theorem hoeffding-iid-bound-ge-general:
  fixes a :: 'a and n :: nat and ε :: real and μ-hat :: real and l u :: real
  assumes a-in: a ∈ A
    and eps-pos: ε ≥ 0
    and bounds: ∀ i < n. ∀ ω ∈ Ω. l ≤ Z i a ω ∧ Z i a ω ≤ u
    and mu-def: μ-hat = (∑ i < n. expectation (λω. Z i a ω))
    and u − l ≠ 0
    and n-pos:  n > 0
    and space-M: space M = Ω
    and sets-M:  sets M  = ℱ
    and indep: indep-vars (λ-. borel) (λi. (λω. Z i a ω)) {i. i < n}
    and rv: ∀ i<n. random-variable borel (λω. Z i a ω)

shows prob {ω ∈ Ω. (∑ i < n. Z i a ω) ≥ μ-hat + ε}
        ≤ exp (− 2 ∗ ε^2 / (real n ∗ (u − l)^2))
proof −
  let ?I = {i. i < n}
  let ?X = λi. (λω. Z i a ω)
  let ?a = λi. l
  let ?b = λi. u

  have finite-I: finite ?I by simp


  have AE-bounds: ∀ i∈?I. AE ω in M. ?a i ≤ ?X i ω ∧ ?X i ω ≤ ?b i
  proof
    fix i assume iI: i ∈ ?I
    have ∀ω∈Ω. l ≤ Z i a ω ∧ Z i a ω ≤ u using bounds iI by simp
    thus AE ω in M. ?a i ≤ ?X i ω ∧ ?X i ω ≤ ?b i
      using assms by auto
  qed


  have indep-loc: indep-interval-bounded-random-variables M ?I ?X ?a ?b
    by (standard; use finite-I indep AE-bounds in auto)

  from indep-loc
  have H: Hoeffding-ineq M ?I ?X ?a ?b
    by (rule Hoeffding-ineq.intro)


  have widths: (∑ i ∈ ?I. (?b i − ?a i)^2) = real n ∗ (u − l)^2
    by simp
  have widths-pos: 0 < (∑ i ∈ ?I. (?b i − ?a i)^2)
    using ‹u − l ≠ 0› n-pos by simp
  have sum-expectations-eq-integrals:
    (∑ i < n. expectation (λω. Z i a ω)) = (∑ i∈?I. integral^L M (?X i))
  proof −
    have eq1: ∀ i<n. expectation (λω. Z i a ω) = integral^L M (?X i)
```

using *rv space-M sets-M* **by** *simp*
moreover have *sum-eq*: $(\sum\ i{<}n.\ integral^L\ M\ (?X\ i)) = (\sum\ i{\in}?I.\ integral^L\ M\ (?X\ i))$
by (*simp add*: *lessThan-def*)
ultimately show *?thesis*
by *simp*
qed

have *sum-integrals-eq*: $(\sum\ i{\in}?I.\ integral^L\ M\ (?X\ i)) = \mu\text{-}hat$
using *mu-def* **by** (*simp add*: *sum-expectations-eq-integrals*)

have *tail*:
prob $\{\omega \in \Omega.\ (\sum\ i{\in}?I.\ ?X\ i\ \omega) - (\sum\ i{\in}?I.\ integral^L\ M\ (?X\ i)) \geq \varepsilon\}$
$\leq exp\ (-\ 2 * \varepsilon\hat{\ }2\ /\ (\sum\ i{\in}?I.\ (?b\ i - ?a\ i)\hat{\ }2))$
using *Hoeffding-ineq.Hoeffding-ineq-ge*[*OF H eps-pos widths-pos*]
by (*smt* (*verit*, *best*) *Collect-cong assms*(*7*))

have *lhs-rewrite*: $\{\omega \in \Omega.\ (\sum\ i < n.\ Z\ i\ a\ \omega) \geq \mu\text{-}hat + \varepsilon\}$
$= \{\omega \in \Omega.\ (\sum\ i{\in}?I.\ ?X\ i\ \omega) - (\sum\ i{\in}?I.\ integral^L\ M\ (?X\ i)) \geq \varepsilon\}$
by (*simp add*: *add.commute le-diff-eq lessThan-def sum-integrals-eq*)

have *rhs-rewrite*: $exp\ (-\ 2 * \varepsilon\hat{\ }2\ /\ (\sum\ i{\in}?I.\ (?b\ i - ?a\ i)\hat{\ }2))$
$= exp\ (-\ 2 * \varepsilon\hat{\ }2\ /\ (real\ n * (u - l)\hat{\ }2))$
using *widths* **by** *simp*

show *?thesis*
using *tail lhs-rewrite rhs-rewrite* **by** *simp*
qed

theorem *hoeffding-iid-bound-le-general*:
fixes $a :: 'a$ **and** $n :: nat$ **and** $\varepsilon :: real$ **and** $\mu\text{-}hat :: real$ **and** $l\ u :: real$
assumes *a-in*: $a \in A$
and *eps-pos*: $\varepsilon \geq 0$
and *bounds*: $\forall i < n.\ \forall \omega \in \Omega.\ l \leq Z\ i\ a\ \omega \wedge Z\ i\ a\ \omega \leq u$
and *mu-def*: $\mu\text{-}hat = (\sum\ i < n.\ expectation\ (\lambda\omega.\ Z\ i\ a\ \omega))$
and $u - l \neq 0$
and *n-pos*: $n > 0$
and *space-M*: $space\ M = \Omega$
and *sets-M*: $sets\ M = \mathcal{F}$
and *indep*: $indep\text{-}vars\ (\lambda\text{-}.\ borel)\ (\lambda i.\ (\lambda\omega.\ Z\ i\ a\ \omega))\ \{i.\ i < n\}$
and *rv*: $\forall\, i{<}n.\ random\text{-}variable\ borel\ (\lambda\omega.\ Z\ i\ a\ \omega)$
shows *prob* $\{\omega \in \Omega.\ (\sum\ i < n.\ Z\ i\ a\ \omega) \leq \mu\text{-}hat - \varepsilon\}$
$\leq exp\ (-\ 2 * \varepsilon\hat{\ }2\ /\ (real\ n * (u - l)\hat{\ }2))$
proof −
let $?I = \{i.\ i < n\}$

```
  let ?X = λi. (λω. Z i a ω)
  let ?a = λi. l
  let ?b = λi. u

  have finite-I: finite ?I by simp


  have AE-bounds: ∀ i∈?I. AE ω in M. ?a i ≤ ?X i ω ∧ ?X i ω ≤ ?b i
  proof
    fix i assume iI: i ∈ ?I
    have ∀ω∈Ω. l ≤ Z i a ω ∧ Z i a ω ≤ u using bounds iI by simp
    thus AE ω in M. ?a i ≤ ?X i ω ∧ ?X i ω ≤ ?b i
      using assms by auto
  qed


  have indep-loc: indep-interval-bounded-random-variables M ?I ?X ?a ?b
    by (standard; use finite-I indep AE-bounds in auto)

  from indep-loc
  have H: Hoeffding-ineq M ?I ?X ?a ?b
    by (rule Hoeffding-ineq.intro)


  have widths: (∑ i ∈ ?I. (?b i − ?a i)^2) = real n ∗ (u − l)^2 by simp
  have widths-pos: 0 < (∑ i ∈ ?I. (?b i − ?a i)^2) using ‹u − l ≠ 0› n-pos by
simp


  have sum-expectations-eq-integrals:
    (∑ i < n. expectation (λω. Z i a ω)) = (∑ i∈?I. integral^L M (?X i))
  proof −
     have eq1: ∀ i<n. expectation (λω. Z i a ω) = integral^L M (?X i) using rv
space-M sets-M by simp
    moreover have sum-eq: (∑ i<n. integral^L M (?X i)) = (∑ i∈?I. integral^L
M (?X i)) by (simp add: lessThan-def)
    ultimately show ?thesis by simp
  qed

  have sum-integrals-eq: (∑ i∈?I. integral^L M (?X i)) = μ-hat
    using mu-def by (simp add: sum-expectations-eq-integrals)


  have tail:
    prob {ω ∈ Ω. (∑ i∈?I. ?X i ω) − (∑ i∈?I. integral^L M (?X i)) ≤ − ε}
     ≤ exp (− 2 ∗ ε^2 / (∑ i∈?I. (?b i − ?a i)^2))
    using Hoeffding-ineq.Hoeffding-ineq-le[OF H eps-pos widths-pos]
    by (smt (verit, best) Collect-cong assms(7))
```

have *lhs-rewrite*: $\{\omega \in \Omega.\ (\sum\ i < n.\ Z\ i\ a\ \omega) \leq \mu\text{-}hat - \varepsilon\}$
$= \{\omega \in \Omega.\ (\sum\ i{\in}?I.\ ?X\ i\ \omega) - (\sum\ i{\in}?I.\ integral^L\ M\ (?X\ i))$
$\leq -\varepsilon\}$
using *add.inverse-inverse*[*of* ε] *add.inverse-inverse*[*of* μ*-hat*] *assms*(*4*)
cancel-ab-semigroup-add-class.diff-right-commute[*of* μ*-hat* μ*-hat* $\sum$ *uuc*<*n*. *Z* *uuc* *a* -]
cancel-ab-semigroup-add-class.diff-right-commute[*of* $-$ μ*-hat* $-$ μ*-hat* ε $-$ μ*-hat*]
cancel-ab-semigroup-add-class.diff-right-commute[*of* ε $-$ μ*-hat* ε $-$ μ*-hat* ε]
cancel-ab-semigroup-add-class.diff-right-commute[*of* ε ε μ*-hat*]
cancel-ab-semigroup-add-class.diff-right-commute[*of* *0* $-$ μ*-hat* ε]
cancel-ab-semigroup-add-class.diff-right-commute[*of* μ*-hat* μ*-hat* ε]
cancel-ab-semigroup-add-class.diff-right-commute[*of* $-$ μ*-hat* $-$ μ*-hat* ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) $-$ μ*-hat*]
cancel-ab-semigroup-add-class.diff-right-commute[*of* ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) $-$ μ*-hat* ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) $-$ μ*-hat*
$\sum$ *uuc*<*n*. *Z* *uuc* *a* -]
cancel-ab-semigroup-add-class.diff-right-commute[*of* $\sum$ *uuc*<*n*. *Z* *uuc* *a* - $\sum$ *uuc*<*n*. *Z* *uuc* *a* - μ*-hat*]
cancel-ab-semigroup-add-class.diff-right-commute[*of* *0* $-$ μ*-hat* $\sum$ *uuc*<*n*. *Z* *uuc* *a* -]
cancel-comm-monoid-add-class.diff-cancel[*of* ε $-$ μ*-hat*] *cancel-comm-monoid-add-class.diff-cancel*[*of* ε]
cancel-comm-monoid-add-class.diff-cancel[*of* μ*-hat*] *cancel-comm-monoid-add-class.diff-cancel*[*of* $-$ μ*-hat*]
cancel-comm-monoid-add-class.diff-cancel[*of* ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) $-$ μ*-hat*]
cancel-comm-monoid-add-class.diff-cancel[*of* $\sum$ *uuc*<*n*. *Z* *uuc* *a* -] *diff-0*
diff-right-mono[*of* ε μ*-hat* $-$ ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) μ*-hat*] *diff-right-mono*[*of* $\sum$ *uuc*<*n*. *Z* *uuc* *a* - μ*-hat* $-$ ε μ*-hat*]
lessThan-def[*of* *n*] *more-arith-simps*(*1*)[*of* $-$ ε ($\sum$ *uuc*<*n*. *Z* *uuc* *a* -) $-$ μ*-hat*]
by *force*

have *rhs-rewrite*: $exp\ (-\ 2 * \varepsilon\hat{}2\ /\ (\sum\ i{\in}?I.\ (?b\ i - ?a\ i)\hat{}2)) = exp\ (-\ 2 * \varepsilon\hat{}2\ /\ (real\ n * (u - l)\hat{}2))$
using *widths* **by** *simp*

show *?thesis* **using** *tail lhs-rewrite rhs-rewrite* **by** *simp*
qed

theorem *hoeffding-iid-ge-delta-bound*:
fixes *a* :: *'a* **and** *n* :: *nat* **and** δ*-hat* :: *real* **and** μ*-hat* :: *real* **and** *l u* :: *real*
assumes *a-in*: $a \in A$
and *delta-bound*: $0 < \delta\text{-}hat\ \delta\text{-}hat \leq 1$
and *bounds*: $\forall\, i{<}n.\ \forall\, \omega{\in}\Omega.\ l \leq Z\ i\ a\ \omega \wedge Z\ i\ a\ \omega \leq u$
and *mu-def*: $\mu\text{-}hat = (\sum i{<}n.\ expectation\ (\lambda\omega.\ Z\ i\ a\ \omega))$
and *n-pos*: $n > 0$
and *eps-pos*: $\varepsilon \geq 0$
and *u-minus-l-nonzero*: $u - l \neq 0$

```
    and space-M: space M = Ω
    and sets-M: sets M = ℱ
    and indep: indep-vars (λ-. borel) (λi. (λω. Z i a ω)) {i. i < n}
    and rv: ∀ i<n. random-variable borel (λω. Z i a ω)
    and eps-expression: ε = sqrt ((real n ∗ (u − l)^2 ∗ ln (1 / δ-hat)) / 2)
  shows prob {ω ∈ Ω. (∑ i<n. Z i a ω) ≥ μ-hat + ε} ≤ δ-hat ∧
        prob {ω ∈ Ω. (∑ i<n. Z i a ω) ≤ μ-hat − ε} ≤ δ-hat
proof −
  have eps-pos: ε ≥ 0
    using assms(7) by auto
  have eps-squared: ε^2 = (real n ∗ (u − l)^2 ∗ ln (1 / δ-hat)) / 2
    using eps-expression by (simp add: assms(3) delta-bound(1))

  have exp-eq: exp (− 2 ∗ ε^2 / (real n ∗ (u − l)^2)) = δ-hat
  proof −
    have − 2 ∗ ε^2 / (real n ∗ (u − l)^2) = − ln (1 / δ-hat)
    proof −
      have ε^2 /( real n ∗ (u − l)^2 ) = ( ln (1 / δ-hat)) / 2
        using assms(6,8) eps-squared by fastforce
      then show ?thesis by linarith
    qed
    also have ... = ln δ-hat
      using delta-bound(1) by (simp add: ln-div)
    finally show ?thesis
      using delta-bound(1) by simp
  qed

  have ge-bound:
    prob {ω ∈ Ω. (∑ i<n. Z i a ω) ≥ μ-hat + ε} ≤ δ-hat
   using hoeffding-iid-bound-ge-general[OF a-in eps-pos bounds mu-def u-minus-l-nonzero

      n-pos space-M sets-M indep rv] exp-eq by simp

  have le-bound:
    prob {ω ∈ Ω. (∑ i<n. Z i a ω) ≤ μ-hat − ε} ≤ δ-hat
   using hoeffding-iid-bound-le-general[OF a-in eps-pos bounds mu-def u-minus-l-nonzero

      n-pos space-M sets-M indep rv] exp-eq by simp

  show ?thesis
    using ge-bound le-bound by simp
qed

lemma add-le-iff:
  fixes x y z :: real
  shows x ≤ y − z ⟷ x − y ≤ −z
  by auto
lemma max-Suc-0-eq-1: max (Suc 0) x = max 1 x
  by simp
```

theorem *ucb-suboptimal-bound-set*:
  fixes t :: *nat*
    and a :: $'a$
    and Δ :: $'a \Rightarrow$ *real*
  assumes *finite-A*: *finite A*
    and *a-in-A*: $a \in A$
    and *a-star-in-A*: *a-star* $\in A$
    and *argmax-exists*: $A \neq \{\}$
    and *subopt-gap*: $\Delta\ a > 0$
    and *a-not-opt*: $\exists\, a'.\ a' \in A \wedge \Delta\ a > 0$
    and *delta-a*: $\Delta\ a = \mu$ *a-star* $- \mu\ a$
    and *ω-in-Ω*: $\omega \in \Omega$
    and *asm*: $\omega \in \{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\}$
    and *setopt*: $\forall\,\omega \in \Omega.\ \exists$ *a-max* $\in A.\ \forall\, b \in A.\ U\ t\ b\ \omega \leq U\ t$ *a-max* ω
    and *A-t-plus-1-maximizes*:
    $\bigwedge t\ \omega\ a.$ *A-t-plus-1* $t\ \omega = a \Longrightarrow a \in A \wedge (\forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega)$
  shows $\{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\} \subseteq$
        $\{\omega \in \Omega.\ U\ t$ *a-star* $\omega \leq \mu$ *a-star*$\} \cup \{\omega \in \Omega.\ \mu$ *a-star* $\leq U\ t\ a\ \omega\}$
proof −
  have *set-result-1*:
    $\{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\} \subseteq \{\omega \in \Omega.\ \forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega\}$
  proof
    fix ω
    assume $\omega \in \{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\}$
    hence *A-t-plus-1* $t\ \omega = a$ **by** *simp*

    from *setopt* **obtain** *a-max* **where**
      a-max $\in A \wedge (\forall\, b \in A.\ U\ t\ b\ \omega \leq U\ t$ *a-max* $\omega)$
      using ‹$\omega \in \{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\}$› **by** *auto*

    hence $\forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega$
      using ‹*A-t-plus-1* $t\ \omega = a$› *A-t-plus-1-maximizes* **by** *auto*

    thus $\omega \in \{\omega \in \Omega.\ \forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega\}$
      using ‹$\omega \in \{\omega \in \Omega.$ *A-t-plus-1* $t\ \omega = a\}$› **by** *blast*
  qed

  have *set-result-2*: $\{\omega \in \Omega.\ \forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega\} \subseteq \{\omega \in \Omega.\ U\ t$ *a-star* ω
$\leq U\ t\ a\ \omega\}$
  proof
    fix ω **assume** *asm*: $\omega \in \{\omega \in \Omega.\ \forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega\}$
    hence $\omega \in \Omega$ **and** *ub*: $\forall\, b \in A.\ U\ t\ a\ \omega \geq U\ t\ b\ \omega$ **by** *simp-all*
    from *a-star-in-A* **have** $U\ t\ a\ \omega \geq U\ t$ *a-star* ω **using** *ub* **by** *simp*

thus $\omega \in \{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le U\ t\ a\ \omega\}$ **using** ‹$\omega \in \Omega$› **by** *simp*
qed

have *set-result-3*: $\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le U\ t\ a\ \omega\} \subseteq$
$\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le \mu\ \textit{a-star}\} \cup \{\omega \in \Omega.\ \mu\ \textit{a-star} \le U\ t\ a\ \omega\}$
proof
fix ω **assume** *asm*: $\omega \in \{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le U\ t\ a\ \omega\}$
hence $\omega \in \Omega$ **and** *le*: $U\ t\ \textit{a-star}\ \omega \le U\ t\ a\ \omega$ **by** *simp-all*
show $\omega \in \{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le \mu\ \textit{a-star}\} \cup \{\omega \in \Omega.\ \mu\ \textit{a-star} \le U\ t\ a\ \omega\}$
proof (*cases* $U\ t\ \textit{a-star}\ \omega \le \mu\ \textit{a-star}$)
case *True*
then show *?thesis* **using** ‹$\omega \in \Omega$› **by** *simp*
next
case *False*
hence $\mu\ \textit{a-star} < U\ t\ \textit{a-star}\ \omega$ **by** *simp*
with *le* **have** $\mu\ \textit{a-star} < U\ t\ a\ \omega$ **by** (*simp add*: *less-le-trans*)
then show *?thesis* **using** ‹$\omega \in \Omega$› **by** *simp*
qed
qed

from *set-result-1 set-result-2 set-result-3* **show** *?thesis* **by** *auto*
qed

theorem *ucb-suboptimal-bound-prob-statement*:
fixes t :: *nat* **and** a :: $'a$ **and** Δ :: $'a \Rightarrow \textit{real}$
assumes *finite-A*: *finite A*
and *a-star-in-A*: $\textit{a-star} \in A$
and *argmax-exists*: $A \neq \{\}$
and *subopt-gap*: $\Delta\ a > 0$
and *a-not-opt*: $\exists\, a'.\ a' \in A \wedge \Delta\ a > 0$
and *ω-in-Ω*: $\omega \in \Omega$
and *asm*: $\omega \in \{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\}$
and *setopt*: $\forall\, \omega \in \Omega.\ \exists\, \textit{a-max} \in A.\ \forall\, b \in A.\ U\ t\ b\ \omega \le U\ t\ \textit{a-max}\ \omega$
and *A-t-plus-1-maximizes*:
$\bigwedge t\ \omega\ a.\ \textit{A-t-plus-1}\ t\ \omega = a \Longrightarrow a \in A \wedge (\forall\, b \in A.\ U\ t\ a\ \omega \ge U\ t\ b\ \omega)$
and *a-in-A*: $a \in A$
and *omega-in*: $\omega \in \Omega$
and *subopt-gap*: $\Delta\ a > 0$
and *delta-a*: $\Delta\ a = \mu\ \textit{a-star} - \mu\ a$
and *H-def*: $H = (2 * \ln\ (1\ /\ \delta))\ /\ (\Delta\ a)\hat{}2$
and *E-def*: $E = \{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ H \le \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}$
and *F-def*: $F = \{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \le \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \le \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}$
and *G-def*: $G = \{\omega \in \Omega.\ \mu\ \textit{a-star} \le U\ t\ a\ \omega\} \cap \{\omega \in \Omega.\ H \le \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}$
and *meas-sets*: $E \in \textit{sets}\ M\ F \in \textit{sets}\ M\ G \in \textit{sets}\ M$
and *prob-inter*: $\textit{prob}\ (F \cap G) \equiv \textit{enn2real}\ (\textit{emeasure}\ M\ (F \cap G))$

shows *prob* $(\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}) \leq$
$\textit{prob}\ (\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \leq \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\})$
$+$
$\textit{prob}\ (\{\omega \in \Omega.\ U\ t\ a\ \omega \geq \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\})$
proof −

have *subset-result*: $E \subseteq F \cup G$
proof −
have *step1*: $\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \subseteq$
$\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \leq \mu\ \textit{a-star}\} \cup \{\omega \in \Omega.\ \mu\ \textit{a-star} \leq U\ t\ a\ \omega\}$
using *ucb-suboptimal-bound-set assms* **by** *blast*

then have *step2*:
$\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\} \subseteq$
$(\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \leq \mu\ \textit{a-star}\} \cup \{\omega \in \Omega.\ \mu\ \textit{a-star} \leq U\ t\ a\ \omega\}) \cap \{\omega \in$
$\Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}$
by *auto*

then have *step3*:
$\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\} \subseteq$
$(\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \leq \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}) \cup$
$(\{\omega \in \Omega.\ \mu\ \textit{a-star} \leq U\ t\ a\ \omega\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\})$
by (*auto simp add*: *set-eq-iff*)

then have *step4*:
$E \subseteq F \cup G$ **using** *assms* **by** *simp*
then show *?thesis*
by (*simp add*: *E-def F-def G-def*)
qed

have *bound*: $\textit{prob}\ E \leq \textit{prob}\ F + \textit{prob}\ G$
proof (*rule union-bound*)
show $E \subseteq F \cup G$ **by** (*simp add*: *subset-result*)
show $E \in \textit{sets}\ M$ **using** *meas-sets* **by** *simp*
show $F \in \textit{sets}\ M$ **using** *meas-sets* **by** *simp*
show $G \in \textit{sets}\ M$ **using** *meas-sets* **by** *simp*
have $\textit{prob}\ (F \cap G) \equiv \textit{enn2real}\ (\textit{emeasure}\ M\ (F \cap G))$ **using** *prob-inter* **by**
simp
have $\textit{prob}\ E \leq \textit{prob}\ (F \cup G)$
using *assms*(*18*,*19*,*20*) *increasingD measure-increasing subset-result* **by** *blast*
have $\textit{prob}\ (F \cup G) = \textit{prob}\ F + \textit{prob}\ G - \textit{prob}\ (F \cap G)$
by (*simp add*: *Int-commute assms*(*19*,*20*) *finite-measure-Diff′ finite-measure-Union′*)
qed

hence *union-bound*:
$\textit{prob}\ (\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}) \leq$
$\textit{prob}\ (\{\omega \in \Omega.\ U\ t\ \textit{a-star}\ \omega \leq \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}) +$
$\textit{prob}\ (\{\omega \in \Omega.\ U\ t\ a\ \omega \geq \mu\ \textit{a-star}\} \cap \{\omega \in \Omega.\ H \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\})$

```
    using bound assms by simp

  show ?thesis
    using union-bound by simp
qed

lemma U-le-μ-pointwise:
  U t a-star ω ≤ μ a-star ⟷
   M-val t a-star ω − μ a-star ≤
  − q ∗ sqrt (ln (1 / δ) / (2 ∗ real (max 1 (N-n t a-star ω))))
  unfolding U-def
  using add-le-iff max-Suc-0-eq-1
  by auto


lemma U-ge-μ-pointwise:
  assumes delta-a: Δ a = μ a-star − μ a
  shows
    U t a ω ≥ μ a-star ⟷
   M-val t a ω − μ a ≥ Δ a − q ∗ sqrt (ln (1 / δ) / (2 ∗ real (max 1 (N-n t a
ω))))
proof −
  have delta-plus-mu-eq:
    Δ a + μ a = μ a-star
  proof −
    have Δ a = μ a-star − μ a
      using delta-a by simp

    hence Δ a + μ a = (μ a-star − μ a) + μ a
      by simp

    also have ... = μ a-star
      by simp

    finally show ?thesis .
  qed
  have U-ge-μ-rewrite:
    U t a ω ≥ μ a-star ⟷ M-val t a ω − μ a ≥ Δ a − q ∗ sqrt (ln (1 / δ) /
(2 ∗ real (max 1 (N-n t a ω))))
    unfolding U-def using add-le-iff max-Suc-0-eq-1 delta-plus-mu-eq by auto
  then show ?thesis by simp
qed

theorem hoeffding-iid--bound-le:
  fixes a :: 'a and n :: nat and ε :: real and μ-hat :: real and l-hat u-hat :: real
    and I :: nat set
    and X-new :: nat ⇒ 'b ⇒ real
    and a-bound b-bound :: nat ⇒ real
  assumes a-in: a ∈ A
```

and b ∈ Ω
and *eps-pos*: ε ≥ 0
and *eps*: ε = abs (u-hat − l-hat) ∗ sqrt (((real n) / 2) ∗ ln (1 / δ))
and δ ≥ 0 ∧ δ ≤ 1
and *t-eq-n*: t = n
and c > 0
and *bounds*: ∀ j < t. ∀ ω ∈ Ω. ∀ a ∈A. l-hat ≤ Z-hat j a ω ∧ Z-hat j a ω ≤ u-hat
and *mu-def*: μ-hat = (∑ j < t. expectation (λω. Z-hat j a ω))
and u-hat − l-hat ≠ 0
and *t-pos*: t > 0
and ∀ t<n. N-n t a-star b > 0
and *n-pos*: n > 0
and M-val t a-star b ≡ (∑ s < t. if π s b = a-star then Z s a-star b else 0) / real (N-n t a-star b)
and *widths*: (∑ i ∈ I. (b-bound i − a-bound i)^2) = (real n) ∗ (u-hat − l-hat)^2
and *space-M*: space M = Ω
and *sets-M*: sets M = ℱ
and *indep*: indep-vars (λ-. borel) (λj. (λω. Z j a ω)) {j. j < t}
and *rv*: ∀ j<t. random-variable borel (λω. Z j a ω)
and ∀ j<t. Z-hat j a-star ω = c ∗ (if π j b = a-star then Z j a-star b else 0)
and indep-interval-bounded-random-variables M I X-new a-bound b-bound
and *indep-loc*: indep-interval-bounded-random-variables M {j. j < t}
(λj. (λω. Z-hat j a-star ω))
(λj. l-hat) (λj. u-hat)
and *H*: Hoeffding-ineq M {j. j < t}
(λj. (λω. Z-hat j a-star ω))
(λj. l-hat) (λj. u-hat)
and *sum-integrals-eq*: (∑ j ∈ {j. j < t}. integral^L M (λω. Z-hat j a-star ω)) = μ-hat
and *rewriting*: prob {ω ∈ Ω. (∑ j < t. Z-hat j a-star ω) − (∑ j < t. expectation (Z-hat j a-star)) ≤ − ε} =
prob {x ∈ space M. (∑ j < t. Z-hat j a-star x) ≤ (∑ j < t. expectation (Z-hat j a-star)) − ε}
shows prob {ω ∈ Ω. (∑ j < n. Z-hat j a-star ω) ≤ μ-hat − ε}
≤ δ
proof −

let ?I = {j. j < t}
let ?X-new = λj. (λω. Z-hat j a-star ω)
let ?a-bound = λj. l-hat
let ?b-bound = λj. u-hat

have
M-val t a-star b ≡ (∑ s < t. if π s b = a-star then Z s a-star b else 0) / real (N-n t a-star b)
using *assms* **by** *simp*

```
  have finite-I: finite ?I by simp


  have AE-bounds: ∀ j ∈ ?I. AE ω in M. ?a-bound j ≤ ?X-new j ω ∧ ?X-new j ω
≤ ?b-bound j
  proof
    fix j assume j ∈ ?I
    then have j < t by simp

     then have bound-j: ∀ ω ∈ Ω. l-hat ≤ Z-hat j a-star ω ∧ Z-hat j a-star ω ≤
u-hat
      using bounds by (simp add: a-in-A)
    then show AE ω in M. ?a-bound j ≤ ?X-new j ω ∧ ?X-new j ω ≤ ?b-bound j
      using space-M sets-M by force
  qed



  have indep-loc: indep-interval-bounded-random-variables M ?I ?X-new ?a-bound
?b-bound
    using assms by simp
  have H: Hoeffding-ineq M ?I ?X-new ?a-bound  ?b-bound
    using assms by simp


  have sum-integrals-eq: (∑ j∈?I. integral^L M (?X-new j)) = μ-hat
    using assms by simp

  have widths:
    (∑ j ∈ ?I. (?b-bound j − ?a-bound j)^2) = (real n) ∗ (u-hat − l-hat)^2
  proof −
    have (∑ j ∈ ?I. (?b-bound j − ?a-bound j)^2) =
        (∑ j ∈ ?I. (u-hat − l-hat)^2)
      by simp
    also have ... = card ?I ∗ (u-hat − l-hat)^2
      by simp
    also have card ?I = t
      by simp
    also have ... = n
      using ‹t = n› assms by blast
    finally show ?thesis
      using assms by fastforce
  qed

  have widths-pos: 0 < (∑ j ∈ ?I. (?b-bound j − ?a-bound j )^2)
    using ‹u-hat − l-hat ≠ 0› widths n-pos by auto

  have res: prob {ω ∈ Ω. (∑ j < t. Z-hat j a-star ω) ≤  (∑ j < t. expectation
```

(*Z-hat j a-star*)) − ε} =
prob {*x* ∈ *space M*. (∑ *j* < *t*. *Z-hat j a-star x*) ≤ (∑ *j* < *t*. *expectation* (*Z-hat j a-star*)) − ε}
using *assms* **by** *simp*

then have *tail*: *prob* {ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) ≤ (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) − ε}
≤ *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*))
using *Hoeffding-ineq.Hoeffding-ineq-le*[*OF H eps-pos widths-pos*]
by (*smt* (*verit*, *best*) *Collect-cong assms*)

have *lhs-rewrite*: {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) ≤ μ*-hat* − ε}
= {ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) − (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) ≤ −ε}
by (*metis* (*mono-tags*, *lifting*) *add-le-iff assms*(*24*,*6*) *lessThan-def*)
have *rhs-rewrite*: *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*)) = *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *widths* **by** *simp*

have *lhs-prob*:*prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) ≤ μ*-hat* − ε}
= *prob* {ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) − (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) ≤ −ε}
using *lhs-rewrite* **by** *simp*

then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) ≤ μ*-hat* − ε} ≤ *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*))
using *lhs-rewrite tail lhs-prob* **by** (*smt* (*verit*, *ccfv-threshold*) *Collect-cong*)
then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) ≤ μ*-hat* − ε} ≤ *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *rhs-rewrite* **by** *linarith*

then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) − μ*-hat*≤ − ε} = *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a-star* ω) − μ*-hat*≤ − *abs* (*u-hat* − *l-hat*) ∗ *sqrt* (((*real n*) / *2*) ∗ *ln* (*1* / δ))}
using *eps* **by** *simp*

have *res1*: *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*)) =
exp (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* /
(*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *eps* **by** *simp*

have *exp* (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* /
(*real n* ∗ (*u-hat* − *l-hat*)^*2*)) =
exp (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *eps eps-pos* **by** *blast*

then have ... =
exp (− *2* ∗ ((*abs* (*u-hat* − *l-hat*)) ^*2* / (*real n* ∗ (*u-hat* − *l-hat*) ^*2*)) ∗ *sqrt* (((*real n* / *2*) ∗ *ln* (*1* / δ))) ^*2*)
by (*metis* (*no-types*, *opaque-lifting*) *more-arith-simps*(*11*) *power-mult-distrib times-divide-eq-left*
times-divide-eq-right)

have ... =
exp (− *2* ∗ ((*abs* (*u-hat* − *l-hat*)) ^*2* / (*real n* ∗ *abs* ((*u-hat* − *l-hat*)) ^*2*)) ∗ *sqrt* (((*real n* / *2*) ∗ *ln* (*1* / δ))) ^*2*)
by *simp*
have ... = *exp* (− *2* ∗ (*1*/ ((*real n*))) ∗ *sqrt* (((*real n* / *2*) ∗ *ln* (*1* / δ))) ^*2*)
using *assms* **by** *simp*

then have ... = *exp* (−*2* ∗ (*1*/ ((*real n*))) ∗ (((*real n* / *2*) ∗ *ln* (*1* / δ))))
using *power2-eq-square* **by** (*smt* (*verit*, *best*)
arithmetic-simps(*51*) *assms*(*10*) *eps eps-pos real-sqrt-ge-0-iff real-sqrt-pow2 zero-le-mult-iff*)

then have ... = *exp* (−*1*∗ *ln* (*1* / δ))
using *assms* **by** (*simp add*: *field-simps*)
then have *fin*: *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*) ^*2*)) = δ
using δ*-pos assms*
by (*metis* (*lifting*)
‹*exp* (− *2* ∗ (*1* / *real n*) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2) = *exp* (− *2* ∗ (*1* / *real n*) ∗ (*real n* / *2* ∗ *ln* (*1* / δ)))›
‹*exp* (− *2* ∗ (|*u-hat* − *l-hat*| ∗ *sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2 / (*real n* ∗ (*u-hat* − *l-hat*)^2)) = *exp* (− *2* ∗ (|*u-hat* − *l-hat*|^2 / (*real n* ∗ (*u-hat* − *l-hat*)^2)) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2)›
‹*exp* (− *2* ∗ (|*u-hat* − *l-hat*|^2 / (*real n* ∗ (*u-hat* − *l-hat*)^2)) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2) = *exp* (− *2* ∗ (|*u-hat* − *l-hat*|^2 / (*real n* ∗ |*u-hat* − *l-hat*|^2)) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2)›
‹*exp* (− *2* ∗ (|*u-hat* − *l-hat*|^2 / (*real n* ∗ |*u-hat* − *l-hat*|^2)) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2) = *exp* (− *2* ∗ (*1* / *real n*) ∗ (*sqrt* (*real n* / *2* ∗ *ln* (*1* / δ)))^2)›
arith-simps(*56*) *exp-ln ln-divide-pos ln-one more-arith-simps*(*10*) *mult-minus1 of-nat-zero-less-power-iff power-0*)

show *?thesis* **using** *assms fin*
‹*prob* {ω ∈ Ω. (∑*j*<*n*. *Z-hat j a-star* ω) ≤ μ*-hat* − ε} ≤ *exp* (− *2* ∗ ε^2 / (*real n* ∗ (*u-hat* − *l-hat*)^2))›
by *presburger*

qed

theorem *hoeffding-iid--bound-ge*:
fixes *a* :: ′*a* **and** *n* :: *nat* **and** ε :: *real* **and** μ*-hat* :: *real* **and** *l-hat u-hat* :: *real*
and *I* :: *nat set*
and *X-new* :: *nat* ⇒ ′*b* ⇒ *real*

```
    and a-bound b-bound :: nat ⇒ real
  assumes a-in: a ∈ A
    and b ∈ Ω
    and eps-pos: ε ≥ 0
    and eps: ε = abs (u-hat − l-hat) ∗ sqrt (((real n) / 2) ∗ ln (1 / δ))
    and δ ≥ 0 ∧ δ ≤ 1
    and t-eq-n: t = n
    and c > 0
    and bounds: ∀ j < t. ∀ ω ∈ Ω. ∀ a ∈A. l-hat ≤ Z-hat j a ω ∧ Z-hat j a ω ≤
u-hat
    and mu-def: μ-hat = (∑ j < t. expectation (λω. Z-hat j a ω))
    and u-hat − l-hat ≠ 0
    and t-pos: t > 0
    and ∀ t<n. N-n t a-star b > 0
    and n-pos: n > 0
    and M-val t a b ≡ (∑ s < t. if π s b = a then Z s a b else 0) /
 real (N-n t a b)
    and widths: (∑ i ∈ I. (b-bound i − a-bound i)^2) = (real n) ∗ (u-hat −
l-hat)^2
    and space-M: space M = Ω
    and sets-M: sets M = ℱ
    and indep: indep-vars (λ-. borel) (λj. (λω. Z j a ω)) {j. j < t}
    and rv: ∀ j<t. random-variable borel (λω. Z j a ω)
    and ∀ j<t. Z-hat j a ω = c ∗ (if π j b = a then Z j a-star b else 0)
    and indep-interval-bounded-random-variables M I X-new a-bound b-bound
    and indep-loc: indep-interval-bounded-random-variables M {j. j < t}
              (λj. (λω. Z-hat j a ω))
              (λj. l-hat) (λj. u-hat)
    and H: Hoeffding-ineq M {j. j < t}
              (λj. (λω. Z-hat j a ω))
              (λj. l-hat) (λj. u-hat)
    and sum-integrals-eq: (∑ j ∈ {j. j < t}. integral^L M (λω. Z-hat j a ω)) =
μ-hat
    and rewriting: prob {ω ∈ Ω. (∑ j < t. Z-hat j a ω) − (∑ j < t. expectation
(Z-hat j a)) ≥ ε} =
      prob {x ∈ space M. (∑ j < t. Z-hat j a x) ≥ (∑ j < t. expectation (Z-hat j
a)) + ε}
  shows prob {ω ∈ Ω. (∑ j < n. Z-hat j a ω) ≥ μ-hat + ε} ≤ δ
proof −

  let ?I = {j. j < t}
  let ?X-new = λj. (λω. Z-hat j a ω)
  let ?a-bound = λj. l-hat
  let ?b-bound = λj. u-hat

  have
    M-val t a b ≡ (∑ s < t. if π s b = a then Z s a b else 0) / real (N-n t a b)
    using assms by simp
```

have *finite-I*: *finite ?I* **by** *simp*

have *AE-bounds*: $\forall j \in$ *?I*. *AE* ω *in M*. *?a-bound j* $\leq$ *?X-new j* ω $\wedge$ *?X-new j* ω $\leq$ *?b-bound j*
proof
  fix *j* **assume** $j \in$ *?I*
  then have $j < t$ **by** *simp*

  then have *bound-j*: $\forall \omega \in \Omega$. *l-hat* $\leq$ *Z-hat j a* ω $\wedge$ *Z-hat j a* ω $\leq$ *u-hat*
    using *bounds* **by** (*simp add*: *a-in-A*)
  then show *AE* ω *in M*. *?a-bound j* $\leq$ *?X-new j* ω $\wedge$ *?X-new j* ω $\leq$ *?b-bound j*
    using *space-M sets-M* **by** *force*
qed

have *indep-loc*: *indep-interval-bounded-random-variables M ?I ?X-new ?a-bound ?b-bound*
  using *assms* **by** *simp*
have *H*: *Hoeffding-ineq M ?I ?X-new ?a-bound ?b-bound*
  using *assms* **by** *simp*

have *sum-integrals-eq*: ($\sum$ $j \in$*?I*. *integral*^{L} *M* (*?X-new j*)) = μ*-hat*
  using *assms* **by** *simp*

have *widths*:
  ($\sum$ $j \in$ *?I*. (*?b-bound j* $-$ *?a-bound j*)^*2*) = (*real n*) $*$ (*u-hat* $-$ *l-hat*)^*2*
proof $-$
  have ($\sum$ $j \in$ *?I*. (*?b-bound j* $-$ *?a-bound j*)^*2*) =
      ($\sum$ $j \in$ *?I*. (*u-hat* $-$ *l-hat*)^*2*)
    by *simp*
  also have ... = *card ?I* $*$ (*u-hat* $-$ *l-hat*)^*2*
    by *simp*
  also have *card ?I* = *t*
    by *simp*
  also have ... = *n*
    using ‹*t* = *n*› *assms* **by** *blast*
  finally show *?thesis*
    using *assms* **by** *fastforce*
qed

have *widths-pos*: *0* < ($\sum$ $j \in$ *?I*. (*?b-bound j* $-$ *?a-bound j*)^*2*)
  using ‹*u-hat* $-$ *l-hat* $\neq$ *0*› *widths n-pos* **by** *auto*

have *res*: *prob* {$\omega \in \Omega$. ($\sum$ $j < t$. *Z-hat j a* ω) $-$ ($\sum$ $j < t$. *expectation* (*Z-hat j a*)) $\geq \varepsilon$} =

prob {*x* ∈ *space M*. (∑ *j* < *t*. *Z-hat j a x*) ≥ (∑ *j* < *t*. *expectation* (*Z-hat j a*)) + ε}
using *assms* **by** *simp*

then have *tail*: *prob* {ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) ≥ (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) + ε}
≤ *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*))
using *Hoeffding-ineq.Hoeffding-ineq-ge*[*OF H eps-pos widths-pos*]
by (*smt* (*verit*, *best*) *Collect-cong assms*)

have *lhs-rewrite*: {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) ≥ μ*-hat* + ε}
= {ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) − (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) ≥ ε}
using *assms* **by** (*metis* (*lifting*) *ext add.commute le-diff-eq lessThan-def*)
have *rhs-rewrite*: *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*)) = *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *widths* **by** *simp*

have *lhs-prob*:*prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) ≥ μ*-hat* + ε}
= *prob*{ω ∈ Ω. (∑ *j*∈*?I*. *?X-new j* ω) − (∑ *j*∈*?I*. *integral*^{L} *M* (*?X-new j*)) ≥ ε}
using *assms lhs-rewrite* **by** *presburger*

then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) ≥ μ*-hat* + ε} ≤ *exp* (− *2* ∗ ε^*2* / (∑ *j*∈*?I*. (*?b-bound j* − *?a-bound j*)^*2*))
using *lhs-rewrite tail lhs-prob* **by** (*smt* (*verit*, *ccfv-threshold*) *Collect-cong*)
then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) ≥ μ*-hat* + ε} ≤ *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *rhs-rewrite* **by** *linarith*

then have *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) − μ*-hat* ≥ ε} = *prob* {ω ∈ Ω. (∑ *j* < *n*. *Z-hat j a* ω) − μ*-hat* ≥ *abs* (*u-hat* − *l-hat*) ∗ *sqrt* (((*real n*) / *2*) ∗ *ln* (*1* / δ))}
using *eps* **by** *simp*

have *res1*: *exp* (− *2* ∗ ε^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*)) =
exp (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* /
(*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *eps* **by** *simp*

have *exp* (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* /
(*real n* ∗ (*u-hat* − *l-hat*)^*2*)) =
exp (− *2* ∗ (*abs* (*u-hat* − *l-hat*) ∗ *sqrt* ((*real n* / *2*) ∗ *ln* (*1* / δ)))^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*))
using *eps eps-pos* **by** *blast*
then have ... = *exp* (− *2* ∗ ((*abs* (*u-hat* − *l-hat*))^*2* / (*real n* ∗ (*u-hat* − *l-hat*)^*2*

)) $*$ *sqrt* (((*real n* / *2*) $*$ *ln* (*1* / δ)))^*2*)
by (*metis* (*no-types*, *opaque-lifting*) *more-arith-simps*(*11*) *power-mult-distrib times-divide-eq-left*
times-divide-eq-right)

also have ... = *exp* ($-$ *2* $*$ ((*abs* (*u-hat* $-$ *l-hat*))^*2* / (*real n* $*$ *abs* ((*u-hat* $-$ *l-hat*))^*2*)) $*$ *sqrt* (((*real n* / *2*) $*$ *ln* (*1* / δ)))^*2*)
by *simp*
have ... = *exp* ($-$ *2* $*$ (*1*/ ((*real n*))) $*$ *sqrt* (((*real n* / *2*) $*$ *ln* (*1* / δ)))^*2*)
using *assms* **by** *simp*

then have ... = *exp* ($-$*2* $*$ (*1*/ ((*real n*))) $*$ (((*real n* / *2*) $*$ *ln* (*1* / δ))))
using *power2-eq-square* **by** (*smt* (*verit*, *best*)
arithmetic-simps(*51*) *assms*(*10*) *eps eps-pos real-sqrt-ge-0-iff real-sqrt-pow2 zero-le-mult-iff*)

then have ... = *exp* ($-$*1*$*$ *ln* (*1* / δ))
using *assms* **by** (*simp add*: *field-simps*)
then have *fin*: *exp* ($-$ *2* $*$ ε^*2* / (*real n* $*$ (*u-hat* $-$ *l-hat*)^*2*)) = δ
using δ*-pos assms*
by (*metis* (*lifting*)
‹*exp* ($-$ *2* $*$ (*1* / *real n*) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2) = *exp* ($-$ *2* $*$ (*1* / *real n*) $*$ (*real n* / *2* $*$ *ln* (*1* / δ)))›
‹*exp* ($-$ *2* $*$ (|*u-hat* $-$ *l-hat*| $*$ *sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2 / (*real n* $*$ (*u-hat* $-$ *l-hat*)^2)) = *exp* ($-$ *2* $*$ (|*u-hat* $-$ *l-hat*|^2 / (*real n* $*$ (*u-hat* $-$ *l-hat*)^2)) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2)›
‹*exp* ($-$ *2* $*$ (|*u-hat* $-$ *l-hat*|^2 / (*real n* $*$ (*u-hat* $-$ *l-hat*)^2)) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2) = *exp* ($-$ *2* $*$ (|*u-hat* $-$ *l-hat*|^2 / (*real n* $*$ |*u-hat* $-$ *l-hat*|^2)) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2)›
‹*exp* ($-$ *2* $*$ (|*u-hat* $-$ *l-hat*|^2 / (*real n* $*$ |*u-hat* $-$ *l-hat*|^2)) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2) = *exp* ($-$ *2* $*$ (*1* / *real n*) $*$ (*sqrt* (*real n* / *2* $*$ *ln* (*1* / δ)))^2)›
diff-0 exp-ln ln-divide-pos ln-one more-arith-simps(*10*) *mult-minus1 of-nat-zero-less-power-iff power-0*)

show *?thesis* **using** *assms fin*
using ‹*prob* {$\omega \in \Omega$. μ*-hat* + ε $\leq$ ($\sum$*j*<*n*. *Z-hat j a* ω)} $\leq$ *exp* ($-$ *2* $*$ ε^2 / (*real n* $*$ (*u-hat* $-$ *l-hat*)^2))›
by *presburger*

qed

end
end
theory *Discrete-UCB-Step3*
imports *Discrete-UCB-Step2*

begin

locale *bandit-policy* = *bandit-problem* + *prob-space* +
fixes
Ω :: *'b set*
and $\mathcal{F}$ :: *'b set set*
and π :: *nat* ⇒ *'b* ⇒ *'a*
and *N-n* :: *nat* ⇒ *'a* ⇒ *'b* ⇒ *nat*
and *Z* :: *nat* ⇒ *'a* ⇒ *'b* ⇒ *real*
and δ :: *real*
and *q* :: *real*
assumes *fin-A*: *finite A*
and $\omega \in \Omega$
and *a-in-A*: $a \in A$
and *measurable-policy*: $\forall\, t.\ \pi\ t \in$ *measurable M* (*count-space A*)
and *N-n-def*: $\forall\, n\ a\ \omega.$ *N-n n a* ω = *card* $\{t \in \{0..<n\}.\ \pi\ (t+1)\ \omega = a\}$
and *count-assm-pointwise*: $\forall\, n\ \omega.$ ($\sum a \in A.$ *real* (*N-n n a* ω)) = *real n*
and *delta-pos*: $0 < \delta$
and *delta-less1*: $\delta < 1$
and *q-pos*: $q > 0$

begin

definition *sample-mean-Z* :: *nat* ⇒ *'a* ⇒ *'b* ⇒ *real* **where**
sample-mean-Z n a ω = (*1* / *real n*) ∗ ($\sum i<n.$ *Z i a* ω)

definition *M-fun* :: *nat* ⇒ *'a* ⇒ *'b* ⇒ *real* **where**
M-fun t a ω = (*if N-n* (*t+1*) *a* ω = *0 then 0*
else ($\sum$ $s < t.$ (*if* π *s* ω = *a then Z s a* ω *else 0*)) / *real* (*N-n t a* ω))

definition *U* :: *nat* ⇒ *'a* ⇒ *'b* ⇒ *real* **where**
U t a ω = *M-fun t a* ω + *q* ∗ *sqrt* (*ln* (*1* / δ) / (*2* ∗ *real* (*max 1* (*N-n t a* ω))))

definition *A-t-plus-1* :: *nat* ⇒ *'b* ⇒ *'a* **where**
A-t-plus-1 t ω = (*SOME a.* $a \in A \wedge (\forall\, a'.\ a' \in A \longrightarrow$ *U t a* $\omega \geq$ *U t a'* ω))

definition *prob-eq-Ex* :: *'b set* ⇒ *bool* **where**
prob-eq-Ex E ≡ *prob E* = *expectation* ($\lambda\omega.$ *indicator E* ω)

theorem *proposition-15-7*:
assumes
a-in-A: $a \in A$
and $\omega \in \Omega$
and *subopt-gap*: $\Delta\ a > 0$
and *a-not-opt*: $\exists\ a' \in A.\ \Delta\ a' > 0$
and *delta-a*: $\forall\ a \in A.\ \Delta\ a = \mu$ *a-star* $- \mu\ a$
and $k \leq n$
and *from-UCB-step1*: $\forall\ a \in A.$ *expectation* ($\lambda\omega.$ *real* (*N-n n a* ω)) ≤

s n + expectation ($\lambda\omega$. ($\sum$ *t = k..<n. if* π *(t+1)* ω *= a* $\wedge$ *s t* $\leq$ *real (N-n t a* ω*) then 1 else 0*))

and *from-UCB-step2*: $\forall$ *a* $\in$ *A.* $\forall$ *t* $\in$ *{k..<n}. prob* ({ω $\in$ Ω. *A-t-plus-1 t* ω *= a*} $\cap$
{ω $\in$ Ω. *N-n t a* ω $\geq$ (*2* $*$ *ln* (*1* / δ)) / (Δ *a*)^*2*}) $\leq$ *2* $*$ δ

and *eps-pos*: ε > *0*

and *t-gt0*: $\forall$ *t* $\in$ *{k..<n}. t > 0*

and *choice-delta*: $\forall$ *t* $\in$ *{k..<n}.* δ = *1* / (*real t powr* ε)

and *s-form*: $\forall$ *a* $\in$ *A.* $\forall$ *u. s u* = (*2* $*$ ε $*$ *ln* (*real u*)) / ((Δ *a*)^*2*)

and *subset-meas*: $\forall$ *a* $\in$ *A.* $\forall$ *t* $\in$ *{k..<n}.* $\forall$ ω $\in$ Ω. {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*)/(Δ *a*)^*2* $\leq$ *N-n t a* ω} $\subseteq$ Ω

and *prob-eq-E-assm*: $\forall$ *a* $\in$ *A.* $\forall$ *t* $\in$ *{k..<n}. prob* {ω. π (*Suc t*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} =
prob ({ω. π (*Suc t*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} $\cap$ *space M*)

and *finiteness*: $\forall$ *t*$\in$*{k..<n}.* $\forall$ *a*$\in$*A. emeasure M* {ω. π (*t+1*) ω = *a* $\wedge$ *2*$*$$\varepsilon$$*$*ln*(*real t*)/($\Delta$ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} < ∞

and *measurable-set*: $\forall$ *t*$\in$*{k..<n}.* $\forall$ *a*$\in$*A.* {ω. π (*t+1*) ω = *a* $\wedge$ *2*$*$$\varepsilon$$*$*ln*(*real t*)/($\Delta$ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} $\in$ *sets M*

and *eq-sets-optimum*:
$\forall$ *a* $\in$ *A.* $\forall$ *t* $\in$ *{k..<n}.* {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} =
{ω $\in$ Ω. *A-t-plus-1 t* ω = *a*} $\cap$ {ω $\in$ Ω. *N-n t a* ω $\geq$ (*2* $*$ ε $*$ *ln* (*real t*)) / (Δ *a*)^*2*}

shows
$\forall$ *a* $\in$ *A. expectation* ($\lambda\omega$. *real* (*N-n n a* ω)) $\leq$ *s n* + ($\sum$ *t = k..<n. 2* / (*real t powr* ε))

proof –

have *def-sn*: $\forall$ *a* $\in$ *A. s n* = (*2* $*$ ε $*$ *ln* (*real n*)) / ((Δ *a*)^*2*)
using *s-form* **by** *simp*

have *def-st*: $\forall$ *a* $\in$ *A. s t* = (*2* $*$ ε $*$ *ln* (*real t*)) / ((Δ *a*)^*2*)
using *s-form* **by** *simp*

then have *expression*: $\forall$ *a* $\in$ *A. expectation* ($\lambda\omega$. *real* (*N-n n a* ω)) $\leq$
(*2* $*$ ε $*$ *ln* (*real n*)) / ((Δ *a*)^*2*) + *expectation* ($\lambda\omega$. ($\sum$ *t = k..<n. if* π (*t+1*) ω = *a* $\wedge$
(*2* $*$ ε $*$ *ln* (*real t*)) / ((Δ *a*)^*2*) $\leq$ *real* (*N-n t a* ω) *then 1 else 0*))
using *assms def-sn def-st* **by** *simp*

have *eq-if-of-bool*:
$\forall$ *a* $\in$ *A. expectation* ($\lambda\omega$. $\sum$ *t = k..<n.*
(*if* π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*))
= *expectation* ($\lambda\omega$. $\sum$ *t = k..<n.*

of-bool (π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)))
by (*simp add*: *of-bool-def*)

have *eq-indic-bool*:
$\forall$ *a* $\in$ *A*. *expectation* ($\lambda\omega$. $\sum$ *t* = *k..<n*.
of-bool (π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)))
= *expectation* ($\lambda\omega$. $\sum$ *t* = *k..<n*.
indicat-real {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} ω)
by (*simp add*: *indicator-def of-bool-def*)

have *expression-1*:
$\forall$ *a* $\in$ *A*. *expectation* ($\lambda\omega$. $\sum$ *t* = *k..<n*.
(*if* π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*)) =
expectation ($\lambda\omega$. $\sum$ *t* = *k..<n*.
indicat-real {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} ω)
proof
fix *a* **assume** *a* $\in$ *A*
from *eq-if-of-bool*[*rule-format*, *OF* ‹*a* $\in$ *A*›]
have *eq1*: *expectation* ($\lambda\omega$. $\sum$ *t* = *k..<n*.
(*if* π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*)) =
expectation ($\lambda\omega$. $\sum$ *t* = *k..<n*.
of-bool (π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)))
by *simp*

from *eq-indic-bool*[*rule-format*, *OF* ‹*a* $\in$ *A*›]
have *eq2*: *expectation* ($\lambda\omega$. $\sum$ *t* = *k..<n*.
of-bool (π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω))) =
expectation ($\lambda\omega$. $\sum$ *t* = *k..<n*.
indicat-real {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} ω)
by *simp*

from *eq1 eq2* **show**
expectation ($\lambda\omega$. $\sum$ *t* = *k..<n*.
(*if* π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω) *then 1 else 0*)) =
expectation ($\lambda\omega$. $\sum$ *t* = *k..<n*.
indicat-real {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)} ω)
by (*rule trans*)
qed

have *res*:$\forall$ *a* $\in$ *A*. $\forall$ *t* $\in$ {*k..<n*}. *prob-eq-Ex* {ω. π (*t+1*) ω = *a* $\wedge$ *2* $*$ ε $*$ *ln* (*real t*) / (Δ *a*)^*2* $\leq$ *real* (*N-n t a* ω)}

using *assms prob-eq-Ex-def* **by** *auto*

have *lin-of-expect-indicators*:
$\forall\ a \in A.\ \forall t \in \{k..<n\}.\ \ prob\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} =$
$expectation\ (\lambda\omega.\ (indicat\text{-}real\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega))$
using *prob-eq-Ex-def res* **by** *simp*

then have *key-result-1*: $\forall\ a \in A.\ (\sum t = k..<n.$
$expectation\ (\lambda\omega.\ (indicat\text{-}real\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega)))$
$= (\sum t = k..<n.\ prob\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\})$
using *lin-of-expect-indicators* **by** *simp*

have *expression-follow-up*: $\forall\ a \in A.$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ (if\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$(\sum\ t = k..<n.\ prob\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\})$
proof −
have *res1*: $\forall\ a \in A.$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ (if\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.$
$of\text{-}bool\ (\pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)))$
using *eq-if-of-bool* **by** *simp*

have *res2*: $\forall\ a \in A.$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ (if\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ indicat\text{-}real\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega)$
using *expression-1* **by** *simp*

have *res3*: $\forall\ a \in A.$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ (if\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$(\sum\ t = k..<n.\ expectation\ (\lambda\omega.\ indicat\text{-}real\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega))$
proof −

have $\forall\ a \in A.$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ (if\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$expectation\ (\lambda\omega.\ \sum\ t = k..<n.\ indicat\text{-}real\ \{\omega.\ \pi\ (t+1)\ \omega = a \land 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{\ }2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega)$
using *res2* **by** *blast*

then have $\forall\ a \in A.$ *expectation* $(\lambda\omega.\ \sum\ t = k..{<}n.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega) =$
$integral^L\ M\ (\lambda\omega.\ \sum\ t = k..{<}n.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega)$
by *simp*
then have *result-intermed*:$\forall\ a \in A.$ *expectation* $(\lambda\omega.\ \sum\ t = k..{<}n.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega) =$
$(\ \sum\ t = k..{<}n.\ integral^L\ M\ (\lambda\omega.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega))$
using *assms integral-sum* **by** *simp*

have *result*: $\forall\ a \in A.$ *expectation* $(\lambda\omega.\ \sum\ t = k..{<}n.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon *$
$ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega) =$
$(\ \sum\ t = k..{<}n.$ *expectation* $(\lambda\omega.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega))$
using *result-intermed* **by** *auto*
have $\forall\ a \in A.$
expectation $(\lambda\omega.\ \sum\ t = k..{<}n.\ (if\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
expectation $(\lambda\omega.\ \sum\ t = k..{<}n.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega)$
using *res2* **by** *simp*
then have *linearity-expectation*: $\forall\ a \in A.$
expectation $(\lambda\omega.\ \sum\ t = k..{<}n.\ (if\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$(\ \sum\ t = k..{<}n.$ *expectation* $(\lambda\omega.$ *indicat-real* $\{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}\ \omega))$
using *result* **by** *simp*

then show *?thesis* **using** *linearity-expectation* **by** *simp*
qed

have *final-linearity*:$\forall\ a \in A.$
expectation $(\lambda\omega.\ \sum\ t = k..{<}n.\ (if\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0)) =$
$(\sum t = k..{<}n.\ prob\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\})$
using *res1 res2 res3 key-result-1* **by** *simp*
then show *?thesis* **using** *final-linearity* **by** *simp*
qed

have *intermed-result*: $\forall\ a \in A.$
expectation $(\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)) \leq$
$(2 * \varepsilon * ln\ (real\ n))\ /\ (\Delta\ a)\hat{}2\ +$
$(\sum\ t = k..{<}n.\ prob\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\})$
using *expression-follow-up expression* **by** *simp*

then have *follow-up-result*: $\forall\ a \in A.\ \forall\ t \in \{k..<n\}$.
prob $\{\omega.\ \pi\ (t+1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} =$
prob $(\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap \{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\})$
using *expression-follow-up expression eq-sets-optimum intermed-result* **by** *simp*

have *next-result-sum-prob*: $\forall\ a \in A$.
$(\sum\ t = k..<n.\ prob\ \{\omega.\ \pi\ (t+1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\}) =$
$(\sum\ t = k..<n.\ prob\ (\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap \{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\}))$
using *assms follow-up-result* **by** *simp*

then have *next-result-fin*: $\forall\ a \in A$.
expectation $(\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)) \leq$
$(2 * \varepsilon * ln\ (real\ n))\ /\ (\Delta\ a)\hat{}2\ +$
$(\sum\ t = k..<n.\ prob\ (\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap \{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\}))$
using *expression-follow-up next-result-sum-prob expression* **by** *fastforce*

have *generalized-bound*:
$\forall\ a \in A.\ \forall\ t \in \{k..<n\}.\ prob\ (\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap$
$\{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\}) \leq 2\ /\ (real\ t\ powr\ \varepsilon)$
proof
fix *a*
assume *a-in*: $a \in A$
show $\forall\ t \in \{k..<n\}.\ prob\ (\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap \{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\}) \leq 2\ /\ (real\ t\ powr\ \varepsilon)$
proof
fix *t*
assume *t-in*: $t \in \{k..<n\}$
have *Hprob*:
prob $(\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap$
$\{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * ln\ (1\ /\ \delta))\ /\ (\Delta\ a)\hat{}2\}) \leq 2 * \delta$
using *from-UCB-step2*[*rule-format*, *OF a-in t-in*] **by** *blast*

have *ln-eq*: $ln\ (1\ /\ \delta) = \varepsilon * ln\ (real\ t)$
using *choice-delta*[*rule-format*, *OF t-in*] *eps-pos* **by** *simp*

hence *threshold-eq*:
$(2 * ln\ (1\ /\ \delta))\ /\ (\Delta\ a)\hat{}2 = (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2$
by *simp*

hence *set-eq*:
$\{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * ln\ (1\ /\ \delta))\ /\ (\Delta\ a)\hat{}2\} =$
$\{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ (\Delta\ a)\hat{}2\}$

```
      by auto

    from Hprob set-eq have
      prob ({ω ∈ Ω. A-t-plus-1 t ω = a} ∩
          {ω ∈ Ω. N-n t a ω ≥ (2 ∗ ε ∗ ln (real t)) / (Δ a)^2}) ≤ 2 ∗ δ
      by simp

    moreover from choice-delta[rule-format, OF t-in] have 2 ∗ δ = 2 / (real t
powr ε)
      by simp

    ultimately show
      prob ({ω ∈ Ω. A-t-plus-1 t ω = a} ∩
         {ω ∈ Ω. N-n t a ω ≥ (2 ∗ ε ∗ ln (real t)) / (Δ a)^2}) ≤ 2 / (real t powr
ε)
      by simp
  qed
qed

have sum-mono-expression:
  ∀ a ∈ A. (∑ t = k..<n. prob ({ω ∈ Ω. A-t-plus-1 t ω = a} ∩
                  {ω ∈ Ω. N-n t a ω ≥ (2 ∗ ε ∗ ln (real t)) / (Δ a)^2}))
  ≤ (∑ t = k..<n. 2 / (real t powr ε))
proof
  fix a
  assume a-in: a ∈ A
  show (∑ t = k..<n. prob ({ω ∈ Ω. A-t-plus-1 t ω = a} ∩ {ω ∈ Ω. N-n t a ω
≥ (2 ∗ ε ∗ ln (real t)) / (Δ a)^2}))
  ≤ (∑ t = k..<n. 2 / (real t powr ε))
    using generalized-bound a-in by (intro sum-mono) auto
qed

have final-result:
  ∀ a ∈ A. expectation (λω. real (N-n n a ω))
      ≤ (2 ∗ ε ∗ ln (real n)) / ((Δ a)^2) + (∑ t = k..<n. 2 / (real t powr ε))
proof
  fix a
  assume a-in: a ∈ A
  from next-result-fin[rule-format, OF a-in]
  have bound1:
    expectation (λω. real (N-n n a ω))
     ≤ (2 ∗ ε ∗ ln (real n)) / ((Δ a)^2)
        + (∑ t = k..<n. prob ({ω ∈ Ω. A-t-plus-1 t ω = a}
                          ∩ {ω ∈ Ω. N-n t a ω ≥ (2 ∗ ε ∗ ln (real t)) / ((Δ
a)^2)}))
    by simp
  moreover from sum-mono-expression[rule-format, OF a-in]
  have bound2:
    (∑ t = k..<n. prob ({ω ∈ Ω. A-t-plus-1 t ω = a}
```

$\cap \{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * \varepsilon * ln\ (real\ t))\ /\ ((\Delta\ a)\hat{}2)\}))$
$\leq (\sum\ t = k..<n.\ 2\ /\ (real\ t\ powr\ \varepsilon))$
by *simp*
ultimately show
expectation $(\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega))$
$\leq (2 * \varepsilon * ln\ (real\ n))\ /\ ((\Delta\ a)\hat{}2) + (\sum\ t = k..<n.\ 2\ /\ (real\ t\ powr\ \varepsilon))$
by *auto*
qed

show *?thesis* **using** *assms final-result* **by** *simp*

qed

theorem *theorem-15-4*:
assumes
a-in-A: $a \in A$
and *finite A* **and** $\forall a \in A.\ integrable\ M\ (\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega))$
and ω*-in-*Ω: $\omega \in \Omega$
and *subopt-gap*: $\forall\ a \in A.\ \Delta\ a > 0$
and *a-not-opt*: $\exists\ a' \in A.\ \Delta\ a' > 0$
and *delta-a*: $\forall\ a \in A.\ \Delta\ a = \mu\ a\text{-}star - \mu\ a$
and $k \leq n$
and *n-count-assm-pointwise*: $(\sum a{\in}A.\ real\ (N\text{-}n\ n\ a\ \omega)) = real\ n$
and *expected-regret-prop-15-1*: *expectation* $(\lambda\omega.\ R\text{-}n\ n\ \omega) = (\sum a{\in}A.\ \Delta\ a *$ *expectation* $(\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)))$
and *from-UCB-step1*: $\forall\ a \in A.$ *expectation* $(\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)) \leq$
$s\ n +$ *expectation* $(\lambda\omega.\ (\sum\ t = k..<n.\ if\ \pi\ (t{+}1)\ \omega = a \wedge s\ t \leq real\ (N\text{-}n\ t\ a\ \omega)\ then\ 1\ else\ 0))$
and *from-UCB-step2*: $\forall\ a \in A.\ \ \forall\ t \in \{k..<n\}.\ prob\ (\{\omega \in \Omega.\ A\text{-}t\text{-}plus\text{-}1\ t\ \omega = a\} \cap$
$\{\omega \in \Omega.\ N\text{-}n\ t\ a\ \omega \geq (2 * ln\ (1\ /\ \delta))\ /\ (\Delta\ a)\hat{}2\}) \leq 2 * \delta$
and *eps-pos*: $\varepsilon > 0$
and *t-gt0*: $\forall\ t \in \{k..<n\}.\ t > 0$
and *choice-delta*: $\forall\ t \in \{k..<n\}.\ \delta = 1\ /\ (real\ t\ powr\ \varepsilon)$
and *s-form*: $\forall\ a \in A.\ \forall\ u.\ s\ u = (2 * \varepsilon * ln\ (real\ u))\ /\ ((\Delta\ a)\hat{}2)$
and *subset-meas*:$\forall\ a \in A.\ \forall t \in \{k..<n\}.\ \forall \omega \in \Omega.\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)/(\Delta\ a)\hat{}2 \leq N\text{-}n\ t\ a\ \omega\} \subseteq \Omega$
and *prob-eq-E-assm*: $\forall\ a \in A.\ \forall t \in \{k..<n\}.\ prob\ \{\omega.\ \pi\ (Suc\ t)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} =$
$prob\ (\{\omega.\ \pi\ (Suc\ t)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} \cap space\ M)$
and *finiteness*: $\forall t{\in}\{k..<n\}.\ \forall a{\in}A.\ emeasure\ M\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2{*}\varepsilon{*}ln(real\ t)/(\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} < \infty$
and *measurable-set*: $\forall t{\in}\{k..<n\}.\ \forall a{\in}A.\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2{*}\varepsilon{*}ln(real\ t)/(\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} \in sets\ M$

and *eq-sets-optimum*:
$\forall\ a \in A.\ \forall\ t \in \{k..<n\}.\ \{\omega.\ \pi\ (t{+}1)\ \omega = a \wedge 2 * \varepsilon * ln\ (real\ t)\ /\ (\Delta\ a)\hat{}2 \leq real\ (N\text{-}n\ t\ a\ \omega)\} =$

$\{\omega \in \Omega.\ \textit{A-t-plus-1}\ t\ \omega = a\} \cap \{\omega \in \Omega.\ \textit{N-n}\ t\ a\ \omega \geq (2 * \varepsilon * \ln$ $(\textit{real}\ t)) / (\Delta\ a)\hat{}2\}$
and *assms-lin-expect*: $\forall\ a \in A.\ \textit{expectation}\ (\lambda\omega.\ \sum\ t = k..<n.$
$(\textit{if}\ \pi\ (t+1)\ \omega = a \wedge 2 * \varepsilon * \ln\ (\textit{real}\ t) / (\Delta\ a)\hat{}2 \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\ \textit{then}$ $1\ \textit{else}\ 0)) =$
$(\sum\ t = k..<n.\ \textit{expectation}\ (\lambda\omega.\ \textit{indicat-real}\ \{\omega.\ \pi\ (t+1)\ \omega = a \wedge 2 * \varepsilon * \ln$ $(\textit{real}\ t) / (\Delta\ a)\hat{}2 \leq \textit{real}\ (\textit{N-n}\ t\ a\ \omega)\}\ \omega))$
and *mono-sum-sets*:
$(\forall a \in A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq \Delta\ a * (s\ n + (\sum\ t = k..<n.$ $2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$
$\Longrightarrow (\sum a \in A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega))) \leq$
$(\sum a \in A.\ \Delta\ a * (s\ n + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$

shows $\textit{expectation}\ (\lambda\omega.\ \textit{R-n}\ n\ \omega) \leq (\sum a{\in}A.\ \Delta\ a * ((2 * \varepsilon * \ln\ (\textit{real}\ n)) /$ $((\Delta\ a)\hat{}2) + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$
proof −
have *exp-regret-eq*:
$\textit{expectation}\ (\lambda\omega.\ \textit{R-n}\ n\ \omega) = (\sum a{\in}A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a$ $\omega)))$
using *assms* **by** *fastforce*

have *result-2*: $\forall\ a \in A.\ \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega))$
$\leq s\ n + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))$
using *assms proposition-15-7* **by** (*smt* (*verit*, *ccfv-threshold*) *a-star-in-A*)

have *intermed-step*: $\forall a \in A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq \Delta\ a *$ $(s\ n + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon)))$
proof
fix a **assume** $a \in A$
then have $\textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq s\ n + (\sum\ t = k..<n.\ 2 /$ $(\textit{real}\ t\ \textit{powr}\ \varepsilon))$
using *result-2* **by** *auto*
moreover have $0 < \Delta\ a$
using *assms subopt-gap* $\langle a \in A\rangle$ **by** *blast*
ultimately show $\Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq \Delta\ a * (s\ n +$ $(\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon)))$
using *mult-left-mono* **by** *simp*
qed

have *intermed-step-2*:$(\sum a{\in}A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega))) \leq$ $(\sum a{\in}A.\ \Delta\ a * (s\ n + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$
proof −
have *res*:$\forall a{\in}A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq \Delta\ a * (s\ n + (\sum$ $t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon)))$
using *intermed-step* **by** *auto*
have *res2*: $(\forall a \in A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega)) \leq \Delta\ a * (s\ n +$ $(\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$
$\Longrightarrow (\sum a \in A.\ \Delta\ a * \textit{expectation}\ (\lambda\omega.\ \textit{real}\ (\textit{N-n}\ n\ a\ \omega))) \leq (\sum a \in A.\ \Delta\ a * (s$ $n + (\sum\ t = k..<n.\ 2 / (\textit{real}\ t\ \textit{powr}\ \varepsilon))))$

using *res assms* **by** *fastforce*
then have *intermed-step*: $(\sum a{\in}A.\ \Delta\ a * expectation\ (\lambda\omega.\ real\ (N\text{-}n\ n\ a\ \omega)))$
$\leq (\sum a{\in}A.\ \Delta\ a * (s\ n + (\sum\ t = k..{<}n.\ 2\ /\ (real\ t\ powr\ \varepsilon))))$
using *res res2* **by** *auto*

then show *?thesis* **using** *intermed-step* **by** *simp*
qed

then have $expectation\ (\lambda\omega.\ R\text{-}n\ n\ \omega) \leq (\sum a{\in}A.\ \Delta\ a * (s\ n + (\sum\ t = k..{<}n.$
$2\ /\ (real\ t\ powr\ \varepsilon))))$
using *assms intermed-step-2* **by** *linarith*

have $\forall\, a \in A.\ s\ n = (2 * \varepsilon * ln\ (real\ n))\ /\ ((\Delta\ a)\hat{}2)$
using *s-form* **by** *auto*

then have $expectation\ (\lambda\omega.\ R\text{-}n\ n\ \omega) \leq (\sum a{\in}A.\ \Delta\ a * ((2 * \varepsilon * ln\ (real\ n))\ /$
$((\Delta\ a)\hat{}2) + (\sum\ t = k..{<}n.\ 2\ /\ (real\ t\ powr\ \varepsilon))))$
using *assms intermed-step-2* **by** (*metis a-star-in-A less-irrefl verit-minus-simplify*(*1*))

then show *?thesis* **by** *simp*

qed

end
end

References

[1] R. Rebeschini. Lecture 15: Stochastic multi-armed bandit problem and algorithms. 2022. Available at https://www.stats.ox.ac.uk/~rebeschi/teaching/AFoL/22/material/lecture15.pdf.