Recap
The Classical Derivation
A Modern Approach
Universes
Objections, your Honor!

Maximum Entropy and Inductive Logic II

Jürgen Landes

Spring School on Inductive Logic
Outline

1 Recap
2 The Classical Derivation
   - Desiderata
   - The Classical Result
3 A Modern Approach
   - Decision Making
   - Justification
4 Universes
5 Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Rational subjective Beliefs

- Finite propositional language $L$
- Variables $v_1, \ldots, v_n$
- Sentences of $L$, $SL$
- No funny business, self-reference, truth predicates, self-fulfilling
Rational subjective Beliefs

- Finite propositional language $L$
- Variables $v_1, \ldots, v_n$
- Sentences of $L$, $SL$
- No funny business, self-reference, truth predicates, self-fulfilling
- Rational subjective Beliefs
- Finite propositional language $L$
- Variables $v_1, \ldots, v_n$
- Sentences of $L$, $SL$
- No funny business, self-reference, truth predicates, self-fulfilling
Rational subjective Beliefs
Finite propositional language $L$
Variables $v_1, \ldots, v_n$
Sentences of $L$, $SL$
No funny business, self-reference, truth predicates, self-fulfilling
Rational subjective Beliefs

Finite propositional language $L$

Variables $v_1, \ldots, v_n$

Sentences of $L$, $SL$

No funny business, self-reference, truth predicates, self-fulfilling
Probabilities

- \( P : SL \rightarrow [0, 1] \)
- Set of probability functions \( \mathbb{P} \)
- Possible world, states

\[
\omega = v_1 \land v_2 \land \neg v_3 \land \ldots \land v_n \land \neg v_n
\]

\[
P(\varphi) = \sum_{\omega \in \Omega, \omega \models \varphi} P(\omega).
\]

A probability function \( P \in \mathbb{P} \) is uniquely determined by its values on possible worlds, \( \langle P(\omega) : \omega \in \Omega \rangle \).
**Probabilities**

- $P : SL \rightarrow [0, 1]$
- Set of probability functions $\mathbb{P}$
- Possible world, states

$$\omega = v_1 \land v_2 \land \neg v_3 \land \ldots \land v_n \land \neg v_n$$

$$P(\varphi) = \sum_{\omega \in \Omega} P(\omega).$$

A probability function $P \in \mathbb{P}$ is uniquely determined by its values on possible worlds, $\langle P(\omega) : \omega \in \Omega \rangle$. 
Probabilities

- \( P : SL \to [0, 1] \)
- Set of probability functions \( \mathbb{P} \)
- Possible world, states

\[
\omega = v_1 \land v_2 \land \neg v_3 \land \ldots \land v_n \land \neg v_n
\]

\[
P(\varphi) = \sum_{\omega \in \Omega \atop \omega \models \varphi} P(\omega).
\]

- A probability function \( P \in \mathbb{P} \) is uniquely determined by its values on possible worlds, \( \langle P(\omega) : \omega \in \Omega \rangle \).

Recap
The Classical Derivation
A Modern Approach
Universes
Objections, your Honor!

Maximum Entropy and Inductive Logic II
**Probabilities**

- $P : SL \rightarrow [0, 1]$
- Set of probability functions $\mathbb{P}$
- Possible world, states

$$\omega = v_1 \land v_2 \land \neg v_3 \land \ldots \land v_n \land \neg v_n$$

$$P(\varphi) = \sum_{\omega \in \Omega} P(\omega).$$

A probability function $P \in \mathbb{P}$ is uniquely determined by its values on possible worlds, $\langle P(\omega) : \omega \in \Omega \rangle$. 
Probabilities

- $P : SL \rightarrow [0, 1]$
- Set of probability functions $\mathbb{P}$
- Possible world, states

$$\omega = v_1 \land v_2 \land \neg v_3 \land \ldots \land v_n \land \neg v_n$$

$$P(\varphi) = \sum_{\omega \in \Omega} P(\omega) \cdot \omega \models \varphi$$

- A probability function $P \in \mathbb{P}$ is uniquely determined by its values on possible worlds, $\langle P(\omega) : \omega \in \Omega \rangle$. 

Jürgen Landes  Maximum Entropy and Inductive Logic II
Inference Processes

- Knowledge $K$ leads to $E \subseteq P$.
- Formally, an inference process is a map from a set of probability functions (here $E$) to the set of probability functions.
- An inference process is a map (or function)
  - Input: $E$
  - Output: $P^+$. 
Inference Processes

- Knowledge $K$ leads to $\mathbb{E} \subseteq \mathbb{P}$.
- Formally, an *inference process* is a map from a set of probability functions (here $\mathbb{E}$) to the set of probability functions.

An inference process is a map (or function)
- Input: $\mathbb{E}$
- Output: $P^+$.
Inference Processes

- Knowledge $K$ leads to $E \subseteq P$.
- Formally, an *inference process* is a map from a set of probability functions (here $E$) to the set of probability functions.
- An inference process is a map (or function)
  - Input: $E$
  - Output: $P^+$.
Inference Processes

- Knowledge $K$ leads to $E \subseteq P$.
- Formally, an *inference process* is a map from a set of probability functions (here $E$) to the set of probability functions.
- An inference process is a map (or function)
  - Input: $E$
  - Output: $P^+$. 

Jürgen Landes

Maximum Entropy and Inductive Logic II
Inference Processes

- Knowledge $K$ leads to $E \subseteq P$.
- Formally, an *inference process* is a map from a set of probability functions (here $E$) to the set of probability functions.
- An inference process is a map (or function)
- Input: $E$
- Output: $P^+$. 

Jürgen Landes

Maximum Entropy and Inductive Logic II
Obviously right

- Adopt the function, $P^\dagger$, which solve this optimisation problem

$$\text{maximise: } - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$$

subject to: $P \in \mathbb{E}$.

- Shannon Entropy: $H(P) = - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$.

- Maximum Entropy Inference Process (MaxEnt):

$$\{P^+\} = \arg \sup_{P \in \mathbb{E}} H(P)$$
Obviously right

- Adopt the function, $P^t$, which solve this optimisation problem

$$\text{maximise: } - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$$

subject to: $P \in \mathbb{E}$.

- Shannon Entropy: $H(P) = - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$.

- Maximum Entropy Inference Process (MaxEnt):

$$\{P^+\} = \text{arg sup}_{P \in \mathbb{E}} H(P)$$
Obviously right

- Adopt the function, $P^\dagger$, which solve this optimisation problem

\[
\text{maximise: } - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))
\]

subject to: $P \in \mathbb{E}$.

- Shannon Entropy: $H(P) = - \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$.

- Maximum Entropy Inference Process (MaxEnt):

\[
\{ P^+ \} = \arg \sup_{P \in \mathbb{E}} H(P)
\]
Desideratum: Internal

- $P^+ \in \mathbb{E}$
- if $\mathbb{E}$ is non-empty, convex and closed.
Desideratum: Internal

- \( P^+ \in E \)
- if \( E \) is non-empty, convex and closed.
Desideratum: Open-mindedness

\[ P^+(\omega) > 0, \text{ if there exists a } P \in \mathbb{E} \text{ such that } P(\omega) > 0. \]
Desideratum: Language Invariance

- $L'$ generated by $v_1, v_2, \ldots, v_n, v_{n+1}$
- and the same knowledge and the same patient? For all $\varphi \in SL$

$$P'(\varphi) = P^+(\varphi)$$

- Repeat argument for even larger languages.
Desideratum: Language Invariance

- $L'$ generated by $v_1, v_2, \ldots, v_n, v_{n+1}$
- and the same knowledge and the same patient? For all $\phi \in SL$

$$P'(\phi) = P^+(\phi)$$

- Repeat argument for even larger languages.
Desideratum: Language Invariance

- $L'$ generated by $v_1, v_2, \ldots, v_n, v_{n+1}$
- and the same knowledge and the same patient? For all $\varphi \in SL$

\[ P'(\varphi) = P^+(\varphi) \]

- Repeat argument for even larger languages.
Outline

1 Recap

2 The Classical Derivation
   • Desiderata
   • The Classical Result

3 A Modern Approach
   • Decision Making
   • Justification

4 Universes

5 Objections, your Honor!
   • Bug or Feature?
   • Issues of Language
   • Risqué
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Renaming

- Names should not matter.
Knowledge entirely irrelevant to the problem in hand can he ignored.

Languages: $L_1$ with variables $v_1, \ldots, v_s$, $L_2$ with variables $v_{s+1}, \ldots, v_n$.

2 Bodies of Knowledge: $K_1$ formulated within $L_1$, $K_2$ formulated within $L_2$.

For all $\varphi \in L_1$ (problem at hand)

$$IP(\{v_1, \ldots, v_n\}, K_1)(\varphi) = IP(\{v_1, \ldots, v_n\}, K_1 \cup K_2)(\varphi).$$
Irrelevance

Knowledge entirely irrelevant to the problem in hand can be ignored.

Languages: $L_1$ with variables $v_1, \ldots, v_s$, $L_2$ with variables $v_{s+1}, \ldots, v_n$.

2 Bodies of Knowledge: $K_1$ formulated within $L_1$, $K_2$ formulated with in $L_2$

For all $\varphi \in L_1$ (problem at hand)

$$IP(\{v_1, \ldots, v_n\}, K_1)(\varphi) = IP(\{v_1, \ldots, v_n\}, K_1 \cup K_2)(\varphi).$$
Knowledge entirely irrelevant to the problem in hand can be ignored.

Languages: $L_1$ with variables $v_1, \ldots, v_s$, $L_2$ with variables $v_{s+1}, \ldots, v_n$.

2 Bodies of Knowledge: $K_1$ formulated within $L_1$, $K_2$ formulated within $L_2$.

For all $\varphi \in L_1$ (problem at hand)

$$IP(\{v_1, \ldots, v_n\}, K_1)(\varphi) = IP(\{v_1, \ldots, v_n\}, K_1 \cup K_2)(\varphi).$$
Irrelevance

- Knowledge entirely irrelevant to the problem in hand can be ignored.
- Languages: $L_1$ with variables $v_1, \ldots, v_s$, $L_2$ with variables $v_{s+1}, \ldots, v_n$.
- 2 Bodies of Knowledge: $K_1$ formulated within $L_1$, $K_2$ formulated with in $L_2$.
- For all $\varphi \in L_1$ (problem at hand)

$$IP(\{v_1, \ldots, v_n\}, K_1)(\varphi) = IP(\{v_1, \ldots, v_n\}, K_1 \cup K_2)(\varphi).$$
Obstinacy

- Learning something one already believes should not make any difference.
- Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)
- If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then

$$IP(K_1) = IP(K_1 \cup K_2).$$
Learning something one already believes should not make any difference.

Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)

If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then

$$IP(K_1) = IP(K_1 \cup K_2).$$
Learning something one already believes should not make any difference.

Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)

If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then

\[ IP(K_1) = IP(K_1 \cup K_2). \]
Obstinacy

- Learning something one already believes should not make any difference.
- Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)
- If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then

$$IP(K_1) = IP(K_1 \cup K_2).$$
Obstinacy

- Learning something one already believes should not make any difference.
- Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)
- If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then
  $$IP(K_1) = IP(K_1 \cup K_2).$$
Obstinacy

- Learning something one already believes should not make any difference.
- Given two consistent bodies of knowledge, $K_1, K_2$ (on the same language)
- If $IP(K_1)$ (which is a probability function), is consistent with $K_2$, then
- $IP(K_1) = IP(K_1 \cup K_2)$. 
Suppose you know the probability of $\varphi$.

That is, for all $P \in \mathbb{E}$ there exists some $c \in [0, 1]$ such that $P(\varphi) = c$.

If $\omega^- \models \varphi$, then $IP(\omega^-)$ should not depend on your knowledge about the $\neg \varphi$-worlds.
Suppose you know the probability of $\varphi$.
That is, for all $P \in \mathbb{E}$ there exists some $c \in [0, 1]$ such that $P(\varphi) = c$.
If $\omega^- \models \varphi$, then $IP(\omega^-)$ should not depend on your knowledge about the $\neg \varphi$-worlds.
Suppose you know the probability of $\varphi$.

That is, for all $P \in \mathbb{P}$ there exists some $c \in [0, 1]$ such that $P(\varphi) = c$.

If $\omega^- \models \varphi$, then $IP(\omega^-)$ should not depend on your knowledge about the $\neg \varphi$-worlds.
Independence

If $K$ does not contain any information which makes $v_1, v_2$ conditionally dependent on $v_3$, then $v_1, v_2$ should be conditionally independent given $v_3$.

If $K = \{ P(v_3) = \gamma, P(v_1|v_3) = \frac{\alpha}{\gamma}, P(v_2|v_3) = \frac{\beta}{\gamma} \}$ ($P(\gamma) > 0$), then

$$IP(K)(v_1 \land v_2|v_3) = \frac{\alpha \beta}{\gamma \gamma}.$$
Independence

- If $K$ does not contain any information which makes $v_1, v_2$ conditionally dependent on $v_3$, then $v_1, v_2$ should be conditionally independent given $v_3$.
- If $K = \{ P(v_3) = \gamma, P(v_1|v_3) = \frac{\alpha}{\gamma}, P(v_2|v_3) = \frac{\beta}{\gamma} \} \ (P(\gamma) > 0)$, then

$$IP(K)(v_1 \land v_2|v_3) = \frac{\alpha \beta}{\gamma \gamma}.$$
If $E'$ can be obtained from $E$ by moving or deforming $E$ a bit, then

$$IP(E') \approx IP(E).$$

For all sentences $\varphi \in SL$: $IP(E')(\varphi) \approx IP(E)(\varphi)$. 
Continuity

- If $E'$ can be obtained from $E$ by moving or deforming $E$ a bit, then
- $IP(E') \approx IP(E)$.
- For all sentences $\varphi \in SL$: $IP(E')(\varphi) \approx IP(E)(\varphi)$. 
Continuity

- If $E'$ can be obtained from $E$ by moving or deforming $E$ a bit, then
- $\text{IP}(E') \approx \text{IP}(E)$.
- For all sentences $\varphi \in SL$: $\text{IP}(E')(\varphi) \approx \text{IP}(E)(\varphi)$. 
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Theorem

If $E$ is closed, convex and non-empty, IP satisfies Renaming, Irrelevance, Obstinance, Relativisation, Independence and Continuity, then IP is MaxEnt. MaxEnt satisfies language invariance and open-mindedness and is internal.
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué

Jürgen Landes
Maximum Entropy and Inductive Logic II
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Continuity - Bristolean

- One considers the decision problem of setting degrees of belief
- and wonders which beliefs are *best*.
- A *utility function* $u$ is used to measure the goodness / badness / utility of a belief function.
- Determine which beliefs have the “best utility”.
- The usual caveats for decision making apply: Non-causal, act-state independence, etc.
- Decision theoretic norm still undetermined.
One considers the decision problem of setting degrees of belief and wonders which beliefs are best. A utility function $u$ is used to measure the goodness / bad-ness / utility of a belief function. Determine which beliefs have the “best utility”. The usual caveats for decision making apply: Non-causal, act-state independence, etc. Decision theoretic norm still undetermined.
Continuity - Bristolean

- One considers the decision problem of setting degrees of belief
- and wonders which beliefs are best.
- A utility function $u$ is used to measure the goodness / badness / utility of a belief function.
- Determine which beliefs have the “best utility”.
- The usual caveats for decision making apply: Non-causal, act-state independence, etc.
- Decision theoretic norm still undetermined.
One considers the decision problem of setting degrees of belief

and wonders which beliefs are best.

A utility function $u$ is used to measure the goodness / badness / utility of a belief function.

Determine which beliefs have the “best utility”.

The usual caveats for decision making apply: Non-causal, act-state independence, etc.

Decision theoretic norm still undetermined.
Continuity - Bristolean

- One considers the decision problem of setting degrees of belief.
- and wonders which beliefs are best.
- A utility function $u$ is used to measure the goodness / bad-ness / utility of a belief function.
- Determine which beliefs have the “best utility”.
- The usual caveats for decision making apply: Non-causal, act-state independence, etc.
- Decision theoretic norm still undetermined.
One considers the decision problem of setting degrees of belief and wonders which beliefs are best.

A *utility function* $u$ is used to measure the goodness / badness / utility of a belief function.

Determine which beliefs have the “best utility”.

The usual caveats for decision making apply: Non-causal, act-state independence, etc.

Decision theoretic norm still undetermined.
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
If $\omega$ is the true world, then the utility is $u(\omega, P^+) = \log(P^+(\omega))$.

Expected utility for $P \in E$ is: $\sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega))$.

If $E$ is closed, convex and non-empty, maximise worst case expected utility:

$$\arg \sup_{P^+ \in \mathbb{P}} \inf_{P \in E} \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) = \{P^+\}.$$ 

The latest rage is to understand $u$ as an accuracy measure.

Measure closeness to the truth.
If $\omega$ is the true world, then the utility is $u(\omega, P^+) = \log(P^+(\omega))$.

Expected utility for $P \in \mathbb{E}$ is: $\sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega))$.

If $\mathbb{E}$ is closed, convex and non-empty, maximise worst case expected utility:

$$\arg \sup_{P^+ \in \mathbb{P}} \inf_{P \in \mathbb{E}} \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) = \{P^+\}.$$ 

The latest rage is to understand $u$ as an accuracy measure.

Measure closeness to the truth.
Logarithmic Utility

- If \( \omega \) is the true world, then the utility is \( u(\omega, P^+) = \log(P^+(\omega)) \).
- Expected utility for \( P \in \mathbb{E} \) is: \( \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) \).
- If \( \mathbb{E} \) is closed, convex and non-empty, maximise worst case expected utility:

  \[
  \arg \sup_{P^+ \in \mathbb{P}} \inf_{P \in \mathbb{E}} \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) = \{P^\dag\}.
  \]

- The latest rage is to understand \( u \) as an accuracy measure.
- Measure closeness to the truth.

Jürgen Landes  Maximum Entropy and Inductive Logic II
Logarithmic Utility

- If $\omega$ is the true world, then the utility is $u(\omega, P^+) = \log(P^+(\omega))$.
- Expected utility for $P \in \mathbb{E}$ is: $\sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega))$.
- If $\mathbb{E}$ is closed, convex and non-empty, maximise worst case expected utility:

$$\arg \sup_{P^+ \in \mathbb{P}} \inf_{P \in \mathbb{E}} \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) = \{P^\dagger\}.$$  

- The latest rage is to understand $u$ as an accuracy measure.
- Measure closeness to the truth.
Logarithmic Utility

- If $\omega$ is the true world, then the utility is $u(\omega, P^+) = \log(P^+(\omega))$.
- Expected utility for $P \in \mathbb{E}$ is: $\sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega))$.
- If $\mathbb{E}$ is closed, convex and non-empty, maximise worst case expected utility:

$$\arg \sup_{P^+ \in \mathbb{P}} \inf_{P \in \mathbb{E}} \sum_{\omega \in \Omega} P(\omega) \log(P^+(\omega)) = \{ P^\dagger \}.$$

- The latest rage is to understand $u$ as an accuracy measure.
- Measure closeness to the truth.
Utility Theory

- If you do have a utility function and a Decision Theoretic Norm,
- then you can show that a particular inference process is optimal with respect to the above.
- Justifications in terms of *common-sense* principles hinge on the common-sensicality of the principles.
- Justifications in terms of utility functions *appear* much more objective.
- However, one has to give a story explaining where the utility function and the DTN come from.
Utility Theory

- If you do have a utility function and a Decision Theoretic Norm,
- then you can show that a particular inference process is optimal with respect to the above.
- Justifications in terms of common-sense principles hinge on the common-sensicality of the principles.
- Justifications in terms of utility functions appear much more objective.
- However, one has to give a story explaining where the utility function and the DTN come from.
Utility Theory

- If you do have a utility function and a Decision Theoretic Norm,
- then you can show that a particular inference process is optimal with respect to the above.
- Justifications in terms of common-sense principles hinge on the common-sensicality of the principles.
- Justifications in terms of utility functions appear much more objective.
- However, one has to give a story explaining where the utility function and the DTN come from.
Utility Theory

- If you do have a utility function and a Decision Theoretic Norm,
- then you can show that a particular inference process is optimal with respect to the above.
- Justifications in terms of *common-sense* principles hinge on the common-sensicality of the principles.
- Justifications in terms of utility functions *appear* much more objective.
- However, one has to give a story explaining where the utility function and the DTN come from.
Utility Theory

- If you do have a utility function and a Decision Theoretic Norm,
- then you can show that a particular inference process is optimal with respect to the above.
- Justifications in terms of *common-sense* principles hinge on the common-sensicality of the principles.
- Justifications in terms of utility functions *appear* much more objective.
- However, one has to give a story explaining where the utility function and the DTN come from.
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Consider an entire population $M$ with $|M|$ members.

- $K$ consists of statements of the form
- 32.2% of patients complain of symptom $\varphi$

$$\left|\{x \in M \mid x \text{ complains about } \varphi\}\right| \approx \frac{32.2}{100}|M|$$

Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.

Then almost all such populations $M$ satisfy

$$\frac{\left|\{x \in M \mid x \text{ complains about } \varphi\}\right|}{|M|} \approx P^*(\varphi).$$
Ensembles, populations

- Consider an entire population $M$ with $|M|$ members.
- $K$ consists of statements of the form
  - 32.2% of patients complain of symptom $\varphi$

\[
|\{x \in M \mid x \text{ complains about } \varphi\}| \approx \frac{32.2}{100} |M|
\]

- Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.
- Then almost all such populations $M$ satisfy

\[
\frac{|\{x \in M \mid x \text{ complains about } \varphi\}|}{|M|} \approx P^\dagger(\varphi)
\]
Ensembles, populations

- Consider an entire population $M$ with $|M|$ members.
- $K$ consists of statements of the form
- $32.2\%$ of patients complain of symptom $\varphi$

$$\frac{|\{x \in M \mid x \text{ complains about } \varphi\}|}{|M|} \approx \frac{32.2}{100} |M|$$

- Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.
- Then almost all such populations $M$ satisfy

$$\frac{|\{x \in M \mid x \text{ complains about } \varphi\}|}{|M|} \approx P^+(\varphi).$$
Ensembles, populations

- Consider an entire population $M$ with $|M|$ members.
- $K$ consists of statements of the form
- 32.2% of patients complain of symptom $\varphi$

\[
|\{x \in M \mid x \text{ complains about } \varphi\}| \approx \frac{32.2}{100} |M|
\]

- Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.
- Then almost all such populations $M$ satisfy

\[
\frac{|\{x \in M \mid x \text{ complains about } \varphi\}|}{|M|} \approx P^\dagger(\varphi)
\]
Ensembles, populations

- Consider an entire population $M$ with $|M|$ members.
- $K$ consists of statements of the form
- 32.2% of patients complain of symptom $\varphi$
- \[
\left| \left\{ x \in M \mid x \text{ complains about } \varphi \right\} \right| \approx \frac{32.2}{100} |M|
\]
- Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.
- Then almost all such populations $M$ satisfy
- \[
\frac{\left| \left\{ x \in M \mid x \text{ complains about } \varphi \right\} \right|}{|M|} \approx P^+(\varphi).
\]
Ensembles, populations

- Consider an entire population $M$ with $|M|$ members.
- $K$ consists of statements of the form
- 32.2% of patients complain of symptom $\varphi$

$$|\{x \in M \mid x \text{ complains about } \varphi\}| \approx \frac{32.2}{100} |M|$$

- Consider all populations $M$ of fixed large size $|M|$ which satisfy the above.
- Then almost all such populations $M$ satisfy

$$\frac{|\{x \in M \mid x \text{ complains about } \varphi\}|}{|M|} \approx P^\dagger(\varphi).$$
Theorem

For all $\delta > 0$ there exists some natural number $M_0$ such that for all sentences $\varphi \in SL$ and all fixed sizes of populations $|M| \geq M_0$ the proportion of populations $M$ of fixed size $|M| \geq M_0$ which satisfy

$$\left| \frac{\left| \{x \in M \mid x \text{ complains about } \varphi \} \right|}{|M|} - P^\dagger(\varphi) \right| < \delta$$

is greater or equal than $1 - \delta$. 

Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Equivocation, come what may

- The least opinionated function is
  \[ P_\parallel(\omega) = \frac{1}{|\Omega|}, \{P_\parallel\} = \arg\sup_{P\in\mathbb{E}} H(P). \]
- If \( P_\parallel \in \mathbb{E} \), then \( P^\dagger = P_\parallel \).
- Entropy strictly decreases along the rays originating from \( P_\parallel \).
- Entropy maximiser different from \( P_\parallel \) face \( P_\parallel \).
- If \( \mathbb{E} = \{P \in \mathbb{P} \mid 0 \leq P(v_1) \leq 0.5\} \), then \( P^\dagger(v_1) = 0.5 \).
- Can this be right?
Equivocation, come what may

- The least opinionated function is
  \[ P_\omega(\omega) = \frac{1}{|\Omega|}, \{P_\omega\} = \arg \sup_{P \in E} H(P). \]

- If \( P_\omega \in E \), then \( P^\dagger = P_\omega \).

- Entropy strictly decreases along the rays originating from \( P_\omega \).

- Entropy maximiser different from \( P_\omega \) face \( P_\omega \).

- If \( E = \{P \in \mathbb{P} \mid 0 \leq P(v_1) \leq 0.5\} \), then \( P^\dagger(v_1) = 0.5 \).

- Can this be right?
Equivocation, come what may

- The least opinionated function is
  \( P_\|= (\omega) = \frac{1}{|\Omega|}, \{ P_\|= \} = \arg \sup_{P \in \mathbb{E}} H(P) \).
- If \( P_\|= \in \mathbb{E} \), then \( P^\dagger = P_\|= \).
- Entropy strictly decreases along the rays originating from \( P_\|= \).
- Entropy maximiser different from \( P_\|= \) face \( P_\|= \).
- If \( \mathbb{E} = \{ P \in \mathbb{P} \mid 0 \leq P(v_1) \leq 0.5 \} \), then \( P^\dagger(v_1) = 0.5 \).
- Can this be right?
Equivocation, come what may

- The least opinionated function is
  \[ P_\perp(\omega) = \frac{1}{|\Omega|}, \{ P_\perp \} = \arg \sup_{P \in \mathcal{E}} H(P). \]
- If \( P_\perp \in \mathcal{E} \), then \( P^\dagger = P_\perp \).
- Entropy strictly decreases along the rays originating from \( P_\perp \).
- Entropy maximiser different from \( P_\perp \) face \( P_\perp \).
- If \( \mathcal{E} = \{ P \in \mathcal{P} \mid 0 \leq P(v_1) \leq 0.5 \} \), then \( P^\dagger(v_1) = 0.5 \).
- Can this be right?
Equivocation, come what may

- The least opinionated function is
  \[ P_{=}(\omega) = \frac{1}{|\Omega|}, \{P_{=}\} = \arg \sup_{P \in \mathbb{P}} H(P). \]
- If \( P_{=} \in \mathbb{P} \), then \( P^\dagger = P_{=} \).
- Entropy strictly decreases along the rays originating from \( P_{=} \).
- Entropy maximiser different from \( P_{=} \) face \( P_{=} \).
- If \( \mathbb{P} = \{P \in \mathbb{P} | 0 \leq P(v_1) \leq 0.5\} \), then \( P^\dagger(v_1) = 0.5 \).
- Can this be right?
Equivocation, come what may

- The least opinionated function is
  \[ P_\perp(\omega) = \frac{1}{|\Omega|}, \{P_\perp\} = \arg\sup_{P \in \mathcal{E}} H(P). \]

- If \( P_\perp \in \mathcal{E} \), then \( P^\dagger = P_\perp \).

- Entropy strictly decreases along the rays originating from \( P_\perp \).

- Entropy maximiser different from \( P_\perp \) face \( P_\perp \).

- If \( \mathcal{E} = \{ P \in \mathcal{P} \mid 0 \leq P(v_1) \leq 0.5 \} \), then \( P^\dagger(v_1) = 0.5 \).

- Can this be right?
Equivocation, come what may

- The least opinionated function is
- \( P_\bot(\omega) = \frac{1}{|\Omega|}, \{P_\bot\} = \arg \sup_{P \in \mathcal{E}} H(P) \).
- If \( P_\bot \in \mathcal{E} \), then \( P^\dagger = P_\bot \).
- Entropy strictly decreases along the rays originating from \( P_\bot \).
- Entropy maximiser different from \( P_\bot \) face \( P_\bot \).
- If \( \mathcal{E} = \{P \in \mathcal{P} | 0 \leq P(v_1) \leq 0.5\} \), then \( P^\dagger(v_1) = 0.5 \).
- Can this be right?
If $E = P$, then $P^\dagger = P_\subseteq$.

But what if you learn this?
Can this be right?
If $E = P$, then $P^\dagger = P\preceq$.

But what if you learn this?

Can this be right?
If $E = P$, then $P^\dagger = P_{=}$.  
But what if you learn this?  
Can this be right?
Intuition?!?

Maximising $-\sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$ is sooo counter-intuitive.

How would you(!) respond to this objection?

If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Intuition?!?

- Maximising $-\sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$ is sooo counter-intuitive.
- How would you(!) respond to this objection?

- If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Maximising \(-\sum_{\omega \in \Omega} P(\omega) \log(P(\omega))\) is sooo counter-intuitive.

How would you(!) respond to this objection?

If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Maximising $- \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$ is sooo counter-intuitive.

How would you(!) respond to this objection?

If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Maximising $- \sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$ is sooo counter-intuitive.

How would you(!) respond to this objection?

If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Intuition?!!

Maximising $-\sum_{\omega \in \Omega} P(\omega) \log(P(\omega))$ is sooo counter-intuitive.

How would you(!) respond to this objection?

If the aim is to reconstruct “rational” human thinking, then MaxEnt fails; I claim.
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
Dependence

- No knowledge, $E = \mathbb{P}$.
- Possible worlds: red and blue.
  - $P^\uparrow(\text{red}) = \frac{1}{2}$. Okay.
- Possible worlds: red and light blue and dark blue.
  - $P^\uparrow(\text{red}) = \frac{1}{3}$. Ohho!
- This phenomenon is called language dependence.
- It is quite common.
Dependence

- No knowledge, $\mathbb{E} = \mathbb{P}$.
- Possible worlds: red and blue.
  - $P^+(\text{red}) = \frac{1}{2}$. Okay.
- Possible worlds: red and light blue and dark blue.
  - $P^+(\text{red}) = \frac{1}{3}$. Ohho!
- This phenomenon is called language dependence.
- It is quite common.
No knowledge, $E = P$.

Possible worlds: red and blue.

$P^+(red) = \frac{1}{2}$. Okay.

Possible worlds: red and light blue and dark blue.

$P^+(red) = \frac{1}{3}$. Ohho!

This phenomenon is called language dependence.

It is quite common.
Dependence

- No knowledge, $E = P$.
- Possible worlds: red and blue.
  - $P^+(\text{red}) = \frac{1}{2}$. Okay.
- Possible worlds: red and light blue and dark blue.
  - $P^+(\text{red}) = \frac{1}{3}$. Ohho!
- This phenomenon is called language dependence.
- It is quite common.
Dependence

- No knowledge, $\mathbb{E} = \mathbb{P}$.
- Possible worlds: red and blue.
- $P^\dagger(red) = \frac{1}{2}$. Okay.
- Possible worlds: red and light blue and dark blue.
- $P^\dagger(red) = \frac{1}{3}$. Ohho!
- This phenomenon is called language dependence.
- It is quite common.
No knowledge, $E = P$.

Possible worlds: red and blue.

$P^+(\text{red}) = \frac{1}{2}$. Okay.

Possible worlds: red and light blue and dark blue.

$P^+(\text{red}) = \frac{1}{3}$. Ohho!

This phenomenon is called language dependence.

It is quite common.
Dependence

- No knowledge, $E = P$.
- Possible worlds: red and blue.
  - $P^\dagger(\text{red}) = \frac{1}{2}$. Okay.
- Possible worlds: red and light blue and dark blue.
  - $P^\dagger(\text{red}) = \frac{1}{3}$. Ohho!
- This phenomenon is called language dependence.
- It is quite common.
Invariance

\[ L = \langle v_1, v_2, \ldots, v_n \rangle, L' = \langle v_1, v_2, \ldots, v_n, v_{n+1} \rangle. \]

- Knowledge only concerns \( v_1, \ldots, v_n \).
- \( \varphi \in SL. \)

\[ IP(L)(\varphi) = IP(L')(\varphi) \]

- Centre of Mass is not language invariant! MaxEnt is language invariant.
- Ay, caramba!
Invariance

- \( L = \langle v_1, v_2, \ldots, v_n \rangle, \quad L' = \langle v_1, v_2, \ldots, v_n, v_{n+1} \rangle. \)
- Knowledge only concerns \( v_1, \ldots, v_n. \)
- \( \varphi \in SL. \)

\[ IP(L)(\varphi) = IP(L')(\varphi) \]

- Centre of Mass is not language invariant! MaxEnt is language invariant.
- Ay, caramba!
Invariance

- \( L = \langle v_1, v_2, \ldots, v_n \rangle, \quad L' = \langle v_1, v_2, \ldots, v_n, v_{n+1} \rangle. \)
- Knowledge only concerns \( v_1, \ldots, v_n. \)
- \( \varphi \in SL. \)

\[
IP(L)(\varphi) = IP(L')(\varphi)
\]

- Centre of Mass is not language invariant! MaxEnt is language invariant.
- Ay, caramba!
Invariance

\[ L = \langle v_1, v_2, \ldots, v_n \rangle, \quad L' = \langle v_1, v_2, \ldots, v_n, v_{n+1} \rangle. \]

Knowledge only concerns \( v_1, \ldots, v_n \).

\( \varphi \in SL. \)

\[ IP(L)(\varphi) = IP(L')(\varphi) \]

Centre of Mass is not language invariant! MaxEnt is language invariant.

\( \text{Ay, caramba!} \)
Invariance

- \( L = \langle v_1, v_2, \ldots, v_n \rangle, \quad L' = \langle v_1, v_2, \ldots, v_n, v_{n+1} \rangle. \)
- Knowledge only concerns \( v_1, \ldots, v_n. \)
- \( \varphi \in SL. \)

\[
IP(L)(\varphi) = IP(L')(\varphi)
\]

- Centre of Mass is not language invariant! MaxEnt is language invariant.
- Ay, caramba!
Outline

1. Recap
2. The Classical Derivation
   - Desiderata
   - The Classical Result
3. A Modern Approach
   - Decision Making
   - Justification
4. Universes
5. Objections, your Honor!
   - Bug or Feature?
   - Issues of Language
   - Risqué
MaxEnt$_W$:

- An agent ought to equivocate (sufficiently) between the basic possibilities that she can express.
- Language now contains the chance functions $\mathbb{P}$.
- Density functions:

$$
C^1_E := \{ f_1 : E \to [0, 1] : \int_{P \in E} f_1(P) P \, dp \in \mathbb{P} \}.
$$

- “Likelihood of a chance function”
- MaxEnt$_W$: Pick density with greatest entropy.
MaxEnt_\mathcal{W}

- MaxEnt_\mathcal{W}: An agent ought to equivocate (sufficiently) between the basic possibilities that she can express.
- Language now contains the chance functions P.
- Density functions:

\[ C_E^1 := \{ f_1 : \mathbb{E} \rightarrow [0, 1] : \int_{P \in \mathbb{E}} f_1(P)P \, dp \in \mathbb{P} \}. \]

- “Likelihood of a chance function”
- MaxEnt_\mathcal{W}: Pick density with greatest entropy
MaxEnt$_W$

- **MaxEnt$_W$**: An agent ought to equivocate (sufficiently) between the basic possibilities that she can express.
- Language now contains the chance functions $P$.
- Density functions:

$$C^1_E := \{ f_1 : E \to [0, 1] : \int_{P \in E} f_1(P)P \, dp \in P \}.$$ 

- “Likelihood of a chance function”
- MaxEnt$_W$ : Pick density with greatest entropy
MaxEnt\textsubscript{W}

- MaxEnt\textsubscript{W}: An agent ought to equivocate (sufficiently) between the basic possibilities that she can express.
- Language now contains the chance functions \( \mathbb{P} \).
- Density functions:

\[
\mathcal{C}_{E}^{1} := \{ f_{1} : E \to [0,1] : \int_{P \in E} f_{1}(P)P \, dp \in \mathbb{P} \}.
\]

- “Likelihood of a chance function”
- MaxEnt\textsubscript{W}: Pick density with greatest entropy
MaxEnt\textsubscript{W}: An agent ought to equivocate (sufficiently) between the basic possibilities that she can express.

Language now contains the chance functions \( P \).

Density functions:

\[
C_{E}^{1} := \{ f_{1} : E \to [0, 1] : \int_{P \in E} f_{1}(P)P \, dp \in P \}.
\]

“Likelihood of a chance function”

MaxEnt\textsubscript{W}: Pick density with greatest entropy
Entropy: $H(f_1) := - \int_{P \in \mathbb{E}} f_1(P) \cdot \log(f_1(P)) \, dp$.

Let $f_1^\dagger$ be the density in $\mathbb{C}_E^1$ with maximal entropy.

Pick a probability function $P^+ : \int_{P \in \mathbb{E}} f_1^\dagger(P) \, PdP$.

$P^+ = P_{CoM}$. 
Entropy: $H(f_1) := - \int_{P \in \mathbb{E}} f_1(P) \cdot \log(f_1(P)) \, dp$.

Let $f_1^\dagger$ be the density in $\mathcal{C}_\mathbb{E}^1$ with maximal entropy.

Pick a probability function $P^+: \int_{P \in \mathbb{E}} f_1^\dagger(P)P \, dP$.

$P^+ = P_{CoM}$.
Level One

- Entropy: \( H(f_1) := - \int_{P \in \mathbb{E}} f_1(P) \cdot \log(f_1(P)) \, dp \).
- Let \( f_1^\dagger \) be the density in \( C_\mathbb{E}^1 \) with maximal entropy.
- Pick a probability function \( P^+ : \int_{P \in \mathbb{E}} f_1^\dagger(P) \, P \, dP \).
- \( P^+ = P_{CoM} \).
Level One

- Entropy: \( H(f_1) := - \int_{P \in \mathbb{E}} f_1(P) \cdot \log(f_1(P)) dp. \)
- Let \( f_1^\dagger \) be the density in \( \mathbb{C}_E^1 \) with maximal entropy.
- Pick a probability function \( P^+ : \int_{P \in \mathbb{E}} f_1^\dagger(P) PdP. \)
- \( P^+ = P_{CoM} \).
Levels

- Density functions for agents with richer languages

\[
C^{n+1}_E := \{ f_{n+1} : E \to [0, 1] : \int_{f_n \in C^n_E} f_{n+1}(f_n) f_n \, df_n \in C^n_E \}.
\]

- Entropy: \( H(f_{n+1}) := -\int_{f_n \in C^n_E} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n. \)

- Let \( f_{n+1}^\dagger \) be the density in \( C^{n+1}_E \) with maximal entropy (it is flat).

- Pick a \( n \) density \( f_n^\dagger : \int_{f \in C^n_E} f_{n+1}^\dagger(f_n) f_n \, df_n. \)

- Eventually this determines some \( P^{++} \in E. \)

- \( P^{++} = P_{CoM}. \)
Levels

- Density functions for agents with richer languages

\[ C_{\mathbb{E}}^{n+1} := \{ f_{n+1} : \mathbb{E} \to [0, 1] : \int_{f_n \in C_{\mathbb{E}}^n} f_{n+1}(f_n) f_n \, df_n \in C_{\mathbb{E}}^n \}. \]

- Entropy: \( H(f_{n+1}) := - \int_{f_n \in C_{\mathbb{E}}^n} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n. \)

- Let \( f_{n+1}^\dagger \) be the density in \( C_{\mathbb{E}}^{n+1} \) with maximal entropy (it is flat).

- Pick a \( n \) density \( f_n^+ : \int_{f \in C_{\mathbb{E}}^n} f_{n+1}^\dagger(f_n) f_n \, df_n. \)

- Eventually this determines some \( P^{++} \in \mathbb{E}. \)

\[ P^{++} = P_{\text{CoM}}. \]
Levels

- Density functions for agents with richer languages

\[ C_{\mathbb{E}}^{n+1} := \{ f_{n+1} : \mathbb{E} \to [0, 1] : \int_{f_n \in C_{\mathbb{E}}^n} f_{n+1}(f_n)f_n \, df_n \in C_{\mathbb{E}}^n \}. \]

- Entropy: \( H(f_{n+1}) := -\int_{f_n \in C_{\mathbb{E}}^n} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n. \)

- Let \( f_{n+1}^\dagger \) be the density in \( C_{\mathbb{E}}^{n+1} \) with maximal entropy (it is flat).

- Pick a \( n \) density \( f_n^{+} : \int_{f \in C_{\mathbb{E}}^n} f_{n+1}^\dagger(f_n)f_n \, df_n. \)

- Eventually this determines some \( P^{++} \in \mathbb{E}. \)

\[ P^{++} = P_{CoM}. \]
Density functions for agents with richer languages

\[ C_{E}^{n+1} := \{ f_{n+1} : E \to [0, 1] : \int_{f_n \in C_{E}^n} f_{n+1}(f_n) f_n \, df_n \in C_{E}^n \}. \]

Entropy: \( H(f_{n+1}) := -\int_{f_n \in C_{E}^n} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n. \)

Let \( f_{n+1}^\dagger \) be the density in \( C_{E}^{n+1} \) with maximal entropy (it is flat).

Pick a \( n \) density \( f_{n}^+ : \int_{f \in C_{E}^n} f_{n+1}^\dagger(f_n) f_n \, df_n. \)

Eventually this determines some \( P^{++} \in E. \)

\( P^{++} = P_{CoM}. \)
Density functions for agents with richer languages

\[ C^{n+1}_E := \{ f_{n+1} : E \to [0, 1] : \int_{f_n \in C^n_E} f_{n+1}(f_n) f_n \, df_n \in C^n_E \} . \]

Entropy: \( H(f_{n+1}) := -\int_{f_n \in C^n_E} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n . \)

Let \( f^\dagger_{n+1} \) be the density in \( C^{n+1}_E \) with maximal entropy (it is flat).

Pick a \( n \) density \( f_n^+ : \int_{f \in C^n_E} f^\dagger_{n+1}(f_n) f_n \, df_n . \)

Eventually this determines some \( P^{++} \in E . \)

\[ P^{++} = P^{CoM} . \]
Levels

- Density functions for agents with richer languages

$$\mathbb{C}^{n+1}_E := \{ f_{n+1} : E \rightarrow [0, 1] : \int_{f_n \in \mathbb{C}^n_E} f_{n+1}(f_n) f_n \, df_n \in \mathbb{C}^n_E \}. $$

- Entropy: $$H(f_{n+1}) := -\int_{f_n \in \mathbb{C}^n_E} f_{n+1}(f_n) \cdot \log(f_{n+1}(f_n)) \, df_n.$$ 

- Let $$f_{n+1}^{\dagger}$$ be the density in $$\mathbb{C}^{n+1}_E$$ with maximal entropy (it is flat).

- Pick a density $$f_n^{\dagger} : \int_{f \in \mathbb{C}^n_E} f_{n+1}^{\dagger}(f_n) f_n \, df_n.$$ 

- Eventually this determines some $$P^{++} \in \mathbb{E}.$$ 

- $$P^{++} = P_{CoM}.$$
Density Invariance

- Equivocating over a density level ($\geq 1$) leads to centre of mass, regardless of the level.
- Equivocating over $\mathbb{P}$ gives MaxEnt!
- The level matters.
- Ay, ay, caramba!
Density Invariance

- Equivocating over a density level \((\geq 1)\) leads to centre of mass, regardless of the level.
- **Density Invariance**
  - Equivocating over \(P\) gives MaxEnt!
  - The level matters.
  - Ay, ay, caramba!
Density Invariance

- Equivocating over a density level ($\geq 1$) leads to centre of mass, regardless of the level.
- Density Invariance
- Equivocating over $\mathbb{P}$ gives MaxEnt!
- The level matters.
- Ay, ay, caramba!
Density Invariance

- Equivocating over a density level $\geq 1$ leads to centre of mass, regardless of the level.
- Density Invariance
- Equivocating over $P$ gives MaxEnt!
- The level matters.
- Ay, ay, caramba!
Density Invariance

- Equivocating over a density level ($\geq 1$) leads to centre of mass, regardless of the level.
- Density Invariance
- Equivocating over $\mathbb{P}$ gives MaxEnt!
- The level matters.
- Ay, ay, caramba!
References I


References II


References IV


That’s it. Thank you! Questions? – Progic Tomorrow
Intuitive human sensations tend to be logarithmic functions of the stimulus. – Jaynes

Savage

Jon’s L1 – L4