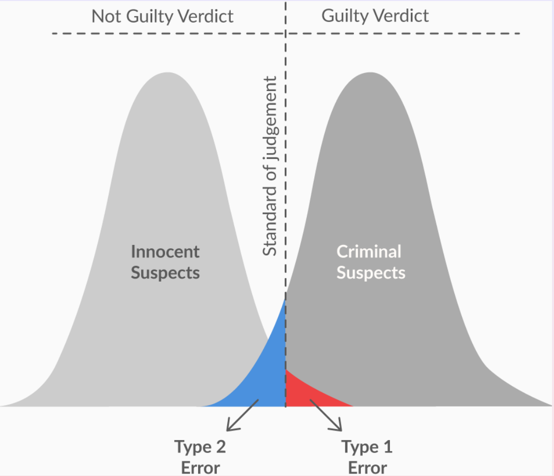
Risky Choice and Blackwell's Theorem

Lones Smith

Madison, 2025

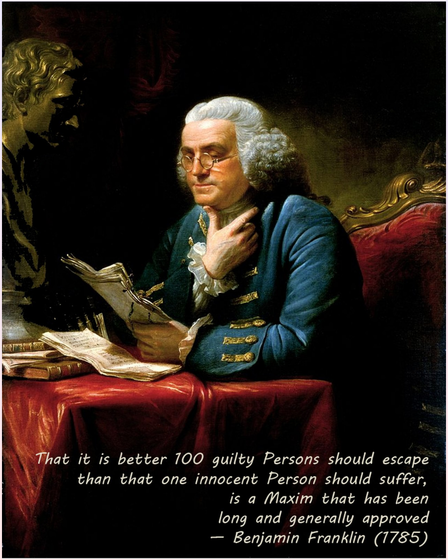
Uncertainty is Key in Guilty & Innocent Verdicts



- The actual ratio of Type I to Type II errors is much smaller than one, in Western legal tradition!

Can We Understand Blackstone's Ratio?

- Blackstone: "Better that ten guilty persons escape, than that one innocent suffer."



Informative Signals

- Two states of the world $\{L, H\}$, and state H has chance q
- Informative signal: a family of probability distributions on observables, *one distribution for each state of the world*
- Seeing σ probabilistically "signals" or indicates the state

	L	H
s	P(s|L)	P(s|H)
t	P(t|L)	P(t|H)

- Here, the realized signal is $\sigma \in \{s, t\}$. Examples:

	L	H
s	2/3	2/3
t	1/3	1/3

uninformative binary signal

	L	H
s	2/3	3/4
t	1/3	1/4

binary signal

	L	H
s	2/3	1/3
t	1/3	2/3

symmetric binary signal

Martingale Property of Beliefs

- *Bayesian updated beliefs are a martingale: After seeing a signal, the expected posterior belief q_1 is the prior q_0.*

$$\begin{aligned} E[q_1|q_0] &= q_0[P(s|H)q_1(s)+P(t|H)q_1(t)] \\ &\quad +(1-q_0)[P(s|L)q_1(s)+P(t|L)q_1(t)] \\ &= q_1(s)[q_0P(s|H)+(1-q_0)P(s|L)] \\ &\quad +q_1(t)[q_0P(t|H)+(1-q_0)P(t|L)] \end{aligned}$$

- Here, we have summed by parts
- By Bayes rule, posterior beliefs are:

$$q_1(s \text{ or } t)=\frac{P(s \text{ or } t|H)q_0}{q_0P(s \text{ or } t|H)+(1-q_0)P(s \text{ or } t|L)}$$

- So $E[q_1|q_0] = q_0P(s|H) + q_0P(t|H) = q_0$
- This is the Law of Iterated Expectations
- Aside: This is a martingale →

Graphical Story of Two State Risky Choice

- *Short* an asset (S), buy it (*long* L), or stay in *cash* (C)
- State $\theta \in \{B, G\}$ fixes *payoffs* ($\theta = G$ with chance q)

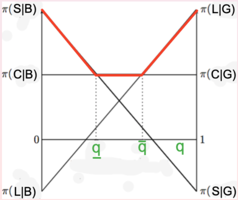
$$\pi(C|G) = \pi(C|B), \quad \pi(L|G) > \pi(L|B), \quad \pi(S|G) < \pi(S|B)$$

- $E(\text{payoff of } a|q) = q\pi(a|G) + (1 - q)\pi(a|B)$ is linear in q

$$\text{Optimal Action is } a^*(q) = \begin{cases} \textit{Short} & \text{if } q \le \underline{q} \\ \textit{Cash} & \text{if } \underline{q} \le q \le \bar{q} \\ \textit{Long} & \text{if } q \ge \bar{q} \end{cases}$$

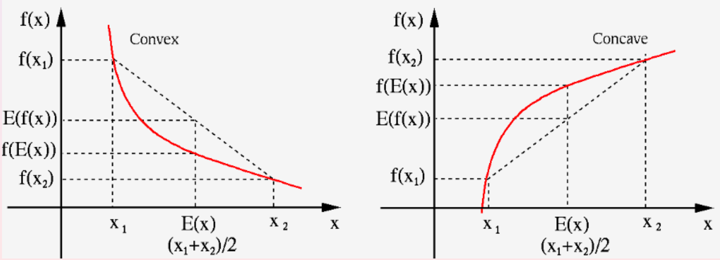
- Fixing $a^*(q)$, payoffs are linear in q
- Optimal payoffs are convex in q if the optimal action changes

expected payoffs

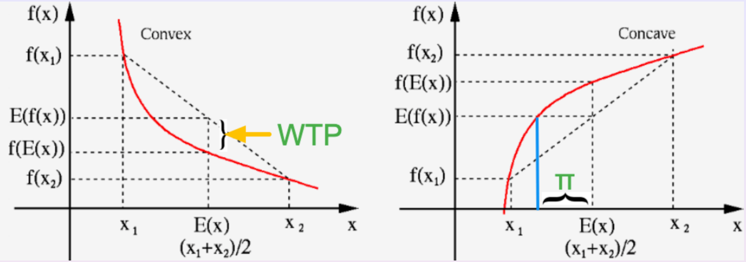


Risk Preference

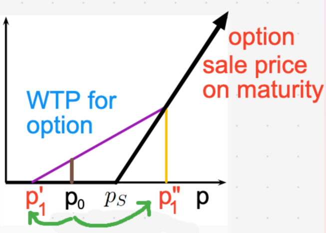
- Risk preference: like/dislike wealth gambles X?
  - risk loving if $\mathbb{E}u(X) \geq u(\mathbb{E}(X))$, sometimes strict
  - risk averse if $\mathbb{E}u(X) \leq u(\mathbb{E}(X))$, sometimes strict
- Jensen's Inequality (1906, Copenhagen Telephone Co!)
  - u is convex on $[a, b]$ iff $u(\mathbb{E}(X)) \geq \mathbb{E}u(X)$ $\forall$ r.v. X on $[a, b]$
  - u is concave on $[a,b]$ iff $u(\mathbb{E}(X)) \leq \mathbb{E}u(X)$ $\forall$ r.v. X on $[a,b]$
  - u is linear on $[a, b]$ iff $u(\mathbb{E}(X)) = \mathbb{E}u(X)$ $\forall$ r,v. X on $[a, b]$



Risk Preference Review

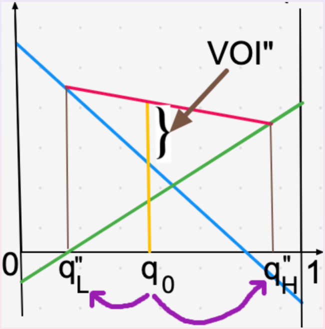


- Concave utility functions: risk premium π measures how much one is WTP to eliminate risk: $u(\mathbb{E}X - \pi) = \mathbb{E}u(X)$
- Induced Convex Payoff Functions
  - E.g. Call Options Induce Risk Loving Behavior by CEOs



(Optionality) Value of Information

- The value of information in a signal σ is the expected optimal payoff given σ minus the prior expected payoff
- E.g.: a binary signal $\Rightarrow$ posterior is $q''_H > q$ or $q''_L < q$



- Claim: The value of information is as depicted.
  - Proof (omitted) uses martingale property of beliefs.
  - So information has zero value if payoffs are locally linear
  - Info has value only if it can change your optimal action
  - It is the value of “optionality”

What is a garbled signal?

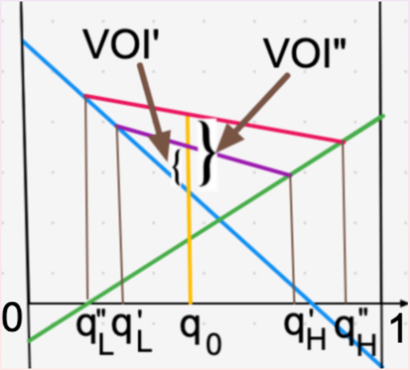
- To get σ' from σ'' by garbling:
  - If signal σ'' gives *t*, send it to *s* with chance $1/6$
  - If signal σ'' gives *s*, send it to *t* with chance $1/6$
  - For instance, in state *H*, the garbling gives *t* with chance

$$(3/4)(5/6) + (1/4)(1/6) = 16/24 = 2/3$$

- The general definition of garbling says that there is a Markov matrix that transforms σ'' into σ'

Baby Blackwell's Theorem (1951)

- Easy two state Bayesian version of Blackwell's Theorem
- **Blackwell (1951)**:
  Garbling a signal reduces the value of information (VOI). Conversely, if the VOI for signal σ'' exceeds that of σ' for all state payoffs, then σ' is a garbling of σ''.
- Blackwell's clever proof used the Minmax Theorem.
- Here's a graphical intuition for easy (⇒) proof:



States and Losses (Payoffs)

- Actual multistate version statistical Blackwell's Theorem
- $\Omega = \{\omega_1, \dots, \omega_n\}$, states of the world
- experiment: n probability measures $(\mu_1, \dots, \mu_n)$ on X
  - Finite outcomes $X = \{x_1, \dots, x_N\}$: an experiment is a Markov matrix of probabilities $P_{n \times N} \equiv [p_{ij}]$, where $\sum_{j=1}^{N} p_{ij} = 1$ and $0 \leq p_{ij} =$ chance of $x_j \in X$ in state ω_i
- $A \subset \mathbb{R}^n$, action space (i.e., vectors of payoffs/losses)
  - $a \in A$ is the n-vector of losses/payoffs in each state, i.e. $a_i =$ loss in state ω_i
- $f : X \to A$, the decision function
  - $f(x_j) \in A$ is the action taken after outcome x_j
- expected loss/payoff from f in state ω_i is $v_i(f)$
  - $v_i(f) \equiv \int_X f_i(x) d\mu_i(x) \equiv \sum_{j=1}^{N} p_{ij} f_i(x_j)$
  - Not Bayesian: We have no prior on Ω
- $B(P, A) \subset \mathbb{R}^n$, loss vector $v(f) = (v_1(f), \dots, v_n(f))$ range

Blackwell's Theorem

- ⋆ $P_{n \times N_1}$ is *more informative* than $Q_{n \times N_2}$ [$P \supset Q$], if
  - any payoff vector attainable with Q is attainable with P
  - $B(P, A) \supseteq B(Q, A)$ *for all* compact convex $A \subset \mathbb{R}^n$.
  - $\Rightarrow$ P has a higher expected value than Q (Baby Blackwell)
- ⋆ Experiment P is *sufficient* for Q [written $P \succ Q$], if
  - i.e. $q_{ij} = \sum_{k=1}^{N_1} p_{ik} m_{kj}$ for all $j = 1, \ldots, N_2$ and $i = 1, \ldots, n$
  - So $PM = Q$ for some Markov matrix $M \leftarrow$ "garbling"

Proposition (Blackwell's Theorem)

$P \succ Q$ iff $P \supset Q$.

- ($\Rightarrow$) is easy: Assume $P \succ Q$.
- $P \supset Q$ if any point in $B(Q, A)$ attainable with a decision function g is attainable under P.
- The decision function $f(x_k) = (\sum_{j=1}^{N_2} m_{kj} g(y_j))$ suffices:
- Why? The payoff under P in state ω_i is

$$v_i(f) = \sum_{k=1}^{N_1} p_{ik} f_i(x_k) = \sum_{k=1}^{N_1} p_{ik} \sum_{j=1}^{N_2} m_{kj} g_i(y_j) = \sum_{j=1}^{N_2} q_{ij} g_i(y_j) = v_i(g)$$

Proof of Hard ($\Leftarrow$) Blackwell's Theorem

- Assume $P \supset Q$.
- $B(P, A) \supseteq B(Q, A)$ $\forall A \subset \mathbb{R}^n$ compact and convex
- Let A be the convex hull of rows of $N_2 \times n$ matrix D
  - i.e. the payoffs in each state after each outcome
- Pick decision function f of (Q, A) picking jth D row for x_j
  - Its expected payoff is $v_i(f) = \sum_{j=1}^{N_2} q_{ij} d_{ji} = (QD)_{ii}$.
- Since $P \supset Q$, some decision function g for (P, A) selects $a^j \in A$ given x_j, with $v_i(g) = \sum_{j=1}^{N_1} p_{ij} a_i^j = v_i(f) \forall i$
- If $a_i^j = \sum_{k=1}^{N_2} m_{jk} d_{ki}$ for a Markov matrix $M \equiv [m_{jk}]$, then PMD and QD have the same diagonal entries:

$$v_i(g) = \sum_{j=1}^{N_1} p_{ij} a_i^j = \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} p_{ij} m_{jk} d_{ki} = (PMD)_{ii}$$

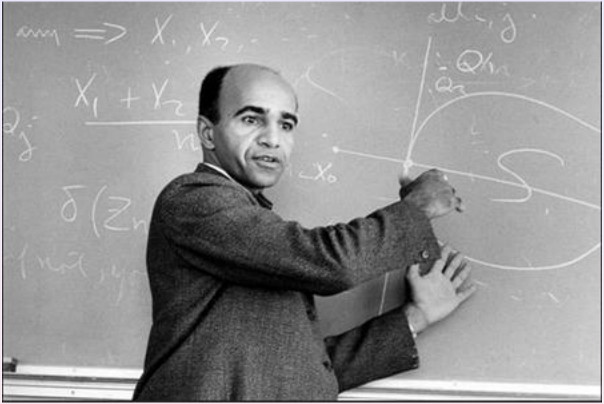
Proof of Hard (⇐) Blackwell's Theorem

- Constant-sum game of decision-maker vs nature.
  - Nature chooses the payoff matrix D and the decision-maker chooses the Markov matrix M.
  - Nature's payoff: $\Pi(D, M) = tr[(PM - Q)D]$
  - *Minimax Theorem* yields a saddle point (D_0, M_0) for the game for all feasible M and D:

$$\Pi(D, M_0) \leq \Pi(D_0, M_0) \leq \Pi(D_0, M)$$

  - He then shows that $PM_0 = Q$, and so $P \succ Q$.

David Blackwell (1919–2010)



- Bottom line: informative signals are rarely ranked — one must be a garbling of the other
- Some pair of decision makers will disagree on a ranking of informative signals
- We next suggest that this conclusion is perhaps too dire