DA5453 — Learning from Human Preferences

Module 1 — Guided take-home assignment

Suryanarayana Sankagiri

Purpose. This ungraded practice assignment will help you develop the principal definitions, calculations, and proof techniques of Module 1. You should attempt the problems in the given order, since later problems may use results established earlier. Your response should include the reasoning, not only the final formula.

Instructions. You should attempt each part independently before consulting references or computational tools. You may discuss questions arising from the assignment in the tutorial and consolidation sessions.

Scope of the assignment and Quiz 1. This assignment is intentionally broader than the material that may ultimately be completed in class. Quiz 1 will be based largely on the questions and ideas appearing here. However, you will be examined only on material covered in lectures, tutorials, or explicitly confirmed during the Module 1 consolidation session.

Disclosure. This handout was generated using a large language model and is constantly reviewed and edited by me, to remove errors or ambiguities. However, some issues may remain. You are requested to report any such issue to me as soon as possible.

Running dataset $\mathcal{D}_0$. Consider four items $\{A, B, C, D\}$. Every unordered pair is compared $m = 10$ times. The observed win counts are as follows.

Pair	Wins by first item	Wins by second item
A, B	$A : 7$	$B : 3$
A, C	$A : 8$	$C : 2$
A, D	$A : 9$	$D : 1$
B, C	$B : 6$	$C : 4$
B, D	$B : 8$	$D : 2$
C, D	$C : 6$	$D : 4$

The dataset $\mathcal{D}_0$ will be used throughout the assignment.

Problem 1 — Preference observations and a cascade model

(a) Suppose that the three outcomes

$$A \succ B, \qquad B \succ C, \qquad C \succ A$$

are observed. Explain why no deterministic ranking of $\{A, B, C\}$ can reproduce all three outcomes. Explain also why these outcomes are not inconsistent with a probabilistic model of comparisons.

(b) Consider the following observations.

(i) An annotator is shown two responses to the same prompt and marks one response as better.

(ii) A customer is shown five laptops and purchases one of them.

(iii) A user is shown a ranked list of documents, scrolls past the first four, and clicks the fifth.

(iv) A reviewer assigns a film a score of four stars out of five.

For each observation, state precisely what has been recorded. Discuss the differences among the four observations. In particular, identify which alternatives were presented, which alternatives were selected or rejected, and what cannot be inferred without additional assumptions about the user or the presentation mechanism.

(c) A ranked list contains items $i_1, i_2, \ldots, i_K$. In the basic *cascade model*, the user examines the items sequentially. On examining item i_k, the user clicks it with probability $a_{i_k} \in (0, 1)$; otherwise the user proceeds to position $k + 1$. The examination ends after the first click. Assume that the click decisions at the examined positions are conditionally independent.

(i) Derive the probability that the first click occurs at position k.

(ii) Derive the probability that no item is clicked.

(iii) Suppose that the first click occurs at position k. Under the stated model, what has been learnt about items $i_1, \ldots, i_{k-1}$? What has been learnt about items $i_{k+1}, \ldots, i_K$?

(iv) Write the likelihood contribution of this observation. Explain why treating the clicked item as having defeated every item below it would not be justified by the cascade model.

Problem 2 — Random utility, MNL, and the IIA property

Let each item i have a systematic utility $\theta_i \in \mathbb{R}$ and realised utility

$$U_i = \theta_i + \varepsilon_i.$$

For a finite choice set S, the selected item is $\arg\max_{i \in S} U_i$. Assume that the noise variables are independent and have the standard Gumbel distribution.

(a) The cumulative distribution function and density of a standard Gumbel random variable are

$$F_\varepsilon(x) = \exp(-e^{-x}), \qquad f_\varepsilon(x) = e^{-x} \exp(-e^{-x}).$$

(i) Verify that $f_\varepsilon = F'_\varepsilon$.

(ii) Write the distribution function $F_i(u) = \Pr(U_i \leq u)$ and density $f_i(u)$ of U_i.

(iii) Explain why ties among the realised utilities occur with probability zero.

(b) Fix $i \in S$. Derive the MNL choice probability in the following steps.

(i) Condition on $U_i = u$ and prove that

$$\Pr(i \text{ is chosen from } S) = \int_{-\infty}^{\infty} f_i(u) \prod_{j \in S \setminus \{i\}} F_j(u)\, du.$$

State explicitly where independence is used.

(ii) Substitute the Gumbel distribution and density into the integral. Show that its integrand can be written as

$$e^{\theta_i} e^{-u} \exp\left\{ -e^{-u} \sum_{j \in S} e^{\theta_j} \right\}.$$

(iii) Evaluate the integral, using an appropriate change of variable, and conclude that

$$\Pr(i \text{ is chosen from } S) = \frac{e^{\theta_i}}{\sum_{j\in S} e^{\theta_j}}.$$

(c) Define $w_i = e^{\theta_i}$.

(i) Rewrite the MNL probability in terms of the positive weights w_i.

(ii) Specialise to $S = \{i, j\}$ and prove

$$\Pr(i \succ j) = \frac{w_i}{w_i + w_j} = \frac{e^{\theta_i}}{e^{\theta_i} + e^{\theta_j}} = \sigma(\theta_i - \theta_j),$$

where $\sigma(x) = 1/(1 + e^{-x})$. Thus identify the BTL model as a special case of MNL.

(iii) Conversely, given arbitrary positive weights (w_i), show how to construct utilities (θ_i) producing the same choice probabilities.

(iv) Explain why the w-parameterisation makes the "share of total weight" interpretation explicit, while the θ-parameterisation makes utility differences and convex optimisation more convenient.

(d) Adding a common constant to every realised utility does not change which item maximises utility.

(i) Use this observation to explain why it is natural that MNL probabilities are unchanged under $\theta_i \mapsto \theta_i + c$.

(ii) Prove the invariance directly from the MNL formula.

(iii) Explain the resulting identifiability issue. Propose a natural restriction that selects one representative from each class of equivalent utility vectors. Give at least two possible restrictions and compare them briefly.

(e) A choice model satisfies *independence of irrelevant alternatives* (IIA) if, for any two items i, j and any two choice sets S, T containing both items,

$$\frac{\Pr(i \mid S)}{\Pr(j \mid S)} = \frac{\Pr(i \mid T)}{\Pr(j \mid T)}.$$

(i) Prove that the MNL model satisfies IIA.

(ii) Let $\theta_A = 2, \theta_B = 1, \theta_C = 0$. Compute $\Pr(A \succ B)$ and $\Pr(A \mid \{A, B, C\})$. Explain why these probabilities need not be equal even though MNL satisfies IIA.

(iii) For the same utilities, compare the likelihood of one three-way choice of A with the likelihood of two independent observations $A \succ B$ and $A \succ C$. Explain why one observation cannot generally be replaced by the other two.

(iv) Give a concrete example in which the IIA property is not plausible.

(f) The following converse is known as *Luce's choice representation.* Let $\mathcal{X}$ be a finite set of items. Suppose that, for every nonempty $S \subseteq \mathcal{X}$, the choice probabilities $p(i \mid S)$ are strictly positive, sum to one, and satisfy IIA. Prove that there exist positive weights (w_i) such that

$$p(i \mid S) = \frac{w_i}{\sum_{j\in S} w_j}.$$

Proceed as follows.

(i) Fix a reference item $0 \in \mathcal{X}$, set $w_0 = 1$, and define

$$w_i = \frac{p(i \mid \{i,0\})}{p(0 \mid \{i,0\})} \quad \text{for } i \neq 0.$$

(ii) For $i, j \neq 0$, apply IIA to the set $\{i, j, 0\}$ and to its two-item subsets to prove that

$$\frac{p(i \mid \{i,j\})}{p(j \mid \{i,j\})} = \frac{w_i}{w_j}.$$

(iii) Apply IIA once more to an arbitrary set S containing i and j. Conclude that $p(i \mid S)/p(j \mid S) = w_i/w_j$.

(iv) Use $\sum_{i \in S} p(i \mid S) = 1$ to obtain the claimed representation.

Problem 3 — The BTL likelihood, gradient, and convexity

For every ordered pair $i \neq j$, let Y_{ij} denote the number of times item i defeats item j, and let $N_{ij} = Y_{ij} + Y_{ji} = N_{ji}$. As established in Problem 2,

$$p_{ij}(\theta) = \Pr(i \succ j) = \sigma(\theta_i - \theta_j), \qquad p_{ji}(\theta) = 1 - p_{ij}(\theta).$$

Comparisons are assumed independent conditional on θ.

(a) By summing over unordered pairs in the form (i, j) with $i < j$, prove that the log-likelihood is

$$\ell(\theta) = \sum_{i<j} \left[Y_{ij} \log p_{ij}(\theta) + Y_{ji} \log p_{ji}(\theta) \right].$$

Write this expression explicitly for the six pairs in $\mathcal{D}_0$. Define the negative log-likelihood by $L(\theta) = -\ell(\theta)$.

(b) Prove that $L(\theta + c\mathbf{1}) = L(\theta)$ for every $c \in \mathbb{R}$. It follows that an unconstrained minimiser cannot be unique. Formulate a constrained optimisation problem that selects a unique representative by imposing, for example, $\sum_i \theta_i = 0$. For this part, assume that a finite minimiser exists and that L is strictly convex under this constraint; you will examine these two conditions in Problem 4.

(c) Prove the gradient formula and then apply it to $\mathcal{D}_0$.

(i) Using $\sigma'(x) = \sigma(x)(1 - \sigma(x))$, differentiate the contribution of one unordered pair and prove that, for every item i,

$$\frac{\partial L}{\partial \theta_i} = \sum_{j \neq i} \left(N_{ij} p_{ij}(\theta) - Y_{ij} \right).$$

Interpret $N_{ij} p_{ij}(\theta) - Y_{ij}$ as "predicted wins minus observed wins".

(ii) At $\theta = (0, 0, 0, 0)$, every $p_{ij} = 1/2$. Compute the total observed wins of each item in $\mathcal{D}_0$, and hence show that

$$\nabla L(0) = (-9, -2, 3, 8)^\top$$

in the order (A, B, C, D).

(iii) Verify that the coordinates of the gradient sum to zero. Explain why this is consistent with shift invariance.

(iv) State which utilities increase and which decrease after one sufficiently small gradient-descent step from the origin. Compare the conclusion with the observed win counts.

(d) Prove that L is a convex function of θ. The following steps may be used.

(i) For one unordered pair (i, j), differentiate twice and show that its Hessian contribution is

$$N_{ij}p_{ij}(\theta)(1 - p_{ij}(\theta))(e_i - e_j)(e_i - e_j)^\top.$$

(ii) For an arbitrary $x \in \mathbb{R}^n$, compute the corresponding quadratic form and show that it is nonnegative.

(iii) Sum the Hessian contributions over all unordered pairs. Conclude that $\nabla^2 L(\theta)$ is positive semidefinite.

(iv) A twice-differentiable function on $\mathbb{R}^n$ is convex if and only if its Hessian is positive semidefinite at every point. Apply this criterion to conclude that $L(\theta)$ is convex.

Problem 4 — Comparison graphs, existence, and uniqueness of the MLE

You will use two graphs associated with the data.

- The *undirected comparison graph* $G = ([n], E)$ contains the edge $\{i, j\}$ when $N_{ij} > 0$. It is *connected* if every two vertices are joined by an undirected path.
- The *directed win graph* $G^{\mathrm{w}} = ([n], E^{\mathrm{w}})$ contains the arc $i \to j$ when $Y_{ij} > 0$. It is *strongly connected* if, for every ordered pair (i, j), there is a directed path from i to j.

(a) Construct both graphs for $\mathcal{D}_0$.

(i) Prove that the undirected comparison graph is connected.

(ii) Prove that the directed win graph is strongly connected.

(iii) After completing parts (d) and (e), return to $\mathcal{D}_0$ and state what the two graph properties imply about existence and uniqueness of the MLE under the constraint $\sum_i \theta_i = 0$.

(b) Consider a dataset $\mathcal{D}_{\text{split}}$ in which comparisons occur only between A and B, and between C and D, with wins in both directions on each of the two observed pairs.

(i) Construct the undirected comparison graph and directed win graph.

(ii) Let θ satisfy $\sum_i \theta_i = 0$. For $t \in \mathbb{R}$, define

$$\theta(t) = \theta + t(1, 1, -1, -1)^\top.$$

Prove that $\sum_i \theta_i(t) = 0$ and $L(\theta(t)) = L(\theta)$ for every t.

(iii) Conclude that the centring constraint alone does not produce a unique MLE for this dataset. Relate this conclusion to the disconnected comparison graph.

(c) Define an alternate dataset $\mathcal{D}_1$ as follows. Every pair is compared, item A defeats each of B, C, D in every comparison, and the comparisons among B, C, D contain wins in both directions.

(i) Construct the undirected comparison graph and directed win graph. Which graph property differs from $\mathcal{D}_0$?

(ii) Consider the centred parameter sequence

$$\theta(t) = (3t, -t, -t, -t)^\top, \qquad t \geq 0.$$

Determine the BTL probabilities involving item A as $t \to \infty$.

(iii) Show that the likelihood contributions from comparisons involving A increase towards their supremum as $t \to \infty$, while no finite value of t attains that supremum. Conclude that $\mathcal{D}_1$ has no finite MLE.

(d) You will now prove how the undirected comparison graph determines the directions in which the loss is strictly convex.

(i) Let
$$a_{ij}(\theta) = N_{ij} p_{ij}(\theta)(1 - p_{ij}(\theta)) \quad \text{for } \{i, j\} \in E.$$
Explain why $a_{ij}(\theta) > 0$ for every finite θ.

(ii) Using Problem 3(d), prove that for every $x \in \mathbb{R}^n$,
$$x^\top \nabla^2 L(\theta) x = \sum_{\{i,j\} \in E} a_{ij}(\theta)(x_i - x_j)^2.$$
A matrix with a quadratic form of this type is called a *weighted graph Laplacian*.

(iii) Prove that x lies in the nullspace of $\nabla^2 L(\theta)$ if and only if x is constant on every connected component of G.

(iv) Deduce that, when G is connected, the nullspace is $\operatorname{span}\{\mathbf{1}\}$ and the Hessian is positive definite on $\{x : \sum_i x_i = 0\}$.

(v) Conclude that, provided a finite MLE exists, connectedness of G makes the centred MLE unique. Explain also why disconnectedness prevents uniqueness in general.

(e) You will now prove how the directed win graph determines whether the centred loss attains its minimum when the undirected comparison graph is connected. First rewrite the negative log-likelihood as
$$L(\theta) = \sum_{i \neq j} Y_{ij} \log\left(1 + e^{\theta_j - \theta_i}\right).$$
Here an ordered pair with $Y_{ij} > 0$ is precisely an arc $i \to j$ in G^{w}.

(i) Assume that G is connected but G^{w} is not strongly connected. Group the vertices of G^{w} into strongly connected components and contract each component to one vertex. The resulting directed graph has no directed cycles. Explain why a finite directed acyclic graph has a source component S, meaning that no arc enters S from outside.

(ii) Since G is connected and S is a proper subset of the vertices, prove that at least one observed edge joins S to its complement. Use the source property to show that every corresponding win arc points from S to its complement.

(iii) Starting from any centred θ, define
$$\theta(t) = \theta + t\mathbf{1}_S - \frac{|S|}{n} t\mathbf{1}, \qquad t \geq 0,$$
where $\mathbf{1}_S$ is the indicator vector of S. Prove that $\theta(t)$ remains centred. Show that loss terms for arcs internal to S or internal to its complement remain unchanged, while every term for an arc from S to its complement decreases with t. Conclude that no finite centred parameter can minimise L.

(iv) Now assume that G^{w} is strongly connected. For a centred vector θ, let
$$R(\theta) = \max_i \theta_i - \min_i \theta_i.$$

Choose a vertex u attaining the minimum and a vertex v attaining the maximum. Strong connectivity gives a directed path $u = i_0 \to i_1 \to \cdots \to i_k = v$ with $k \le n-1$. Prove that at least one arc on this path satisfies

$$\theta_{i_{r+1}} - \theta_{i_r} \ge \frac{R(\theta)}{n-1}.$$

(v) For the arc found above, use $\log(1 + e^x) \ge x$ for $x \ge 0$ and $Y_{i_r i_{r+1}} \ge 1$ to prove

$$L(\theta) \ge \frac{R(\theta)}{n-1}.$$

Show that, on the centred subspace, $\|\theta\|_2 \to \infty$ implies $R(\theta) \to \infty$. Hence $L(\theta) \to \infty$ whenever a centred parameter vector escapes to infinity.

(vi) Use continuity of L and the fact that a closed and bounded set in $\mathbb{R}^{n-1}$ is compact to prove that L attains a finite minimum on the centred subspace. Combine this with part (d) to prove the following statement:

If the undirected comparison graph is connected, the centred BTL MLE exists and is unique if and only if the directed win graph is strongly connected.

Problem 5 — Rank Centrality on $\mathcal{D}_0$

Use the Rank Centrality construction introduced in class. For an observed edge $\{i, j\}$, define the empirical win probability

$$\widehat{p}_{ij} = \frac{Y_{ij}}{N_{ij}}.$$

Choose d at least as large as the maximum degree of the comparison graph, and define a transition matrix P by

$$P_{ij} = \frac{1}{d}\widehat{p}_{ji} \quad (i \ne j,\ \{i, j\} \in E), \qquad P_{ii} = 1 - \sum_{j \ne i} P_{ij}.$$

Thus, the Markov chain tends to move from an item towards items that defeat it.

(a) A probability vector π is a *stationary distribution* of P if

$$\pi^\top P = \pi^\top, \qquad \pi_i \ge 0, \qquad \sum_i \pi_i = 1.$$

(i) Explain why each row of P sums to one.

(ii) Explain why the choice $d \ge d_{\max}$ guarantees $P_{ii} \ge 0$.

(iii) You may use without proof the following Markov-chain result: a finite, irreducible, and aperiodic Markov chain has a unique stationary distribution. State graph and diagonal conditions on P that guarantee irreducibility and aperiodicity.

(b) Construct the empirical transition matrix for $\mathcal{D}_0$.

(i) Take $d = 4$ and use the order (A, B, C, D). Show that

$$P = \frac{1}{40}\begin{pmatrix} 34 & 3 & 2 & 1 \\ 7 & 27 & 4 & 2 \\ 8 & 6 & 22 & 4 \\ 9 & 8 & 6 & 17 \end{pmatrix}.$$

Verify every off-diagonal entry and every row sum.

(ii) Set up the linear equations defining the stationary distribution. Solve them exactly or numerically, and report the ordering of $\pi_A, \pi_B, \pi_C, \pi_D$.

(iii) Explain why a larger stationary probability is interpreted as a larger estimated BTL weight.

(c) You will next justify the method at the population level, using the same four items. Suppose the true BTL weights are $w_A, w_B, w_C, w_D > 0$, and replace the empirical win probabilities in P by their population values:

$$P_{ij} = \frac{1}{d}\frac{w_j}{w_i + w_j}, \qquad i \neq j.$$

(i) Define $\pi_i = w_i / \sum_k w_k$.

(ii) Prove the detailed-balance identity

$$\pi_i P_{ij} = \pi_j P_{ji} \qquad \text{for every } i \neq j.$$

(iii) Sum the detailed-balance identities appropriately to prove $\pi^\top P = \pi^\top$.

(iv) Conclude that population Rank Centrality recovers the BTL weights up to their common scale.

(d) Explain the empirical version of the preceding argument.

(i) For fixed items, what happens to $\widehat{p}_{ij}$ as the number of independent comparisons of each pair tends to infinity?

(ii) Consequently, what should happen to the empirical transition matrix and its stationary distribution?

(iii) Contrast the two estimators studied in this assignment: the MLE directly minimises a negative log-likelihood, whereas Rank Centrality computes a stationary distribution. Explain why they need not agree exactly for a finite dataset even though both target the same population weights.

Problem 6 — A finite-sample recovery bound for the BTL MLE

In this problem, you will derive a recovery theorem for the BTL MLE. For simplicity, you will assume that the comparison graph is complete and that every pair is compared the same number of times. Define the centred, bounded parameter set

$$\Theta_b = \left\{ \theta \in \mathbb{R}^n : \sum_{i=1}^{n} \theta_i = 0,\ \max_i \theta_i - \min_i \theta_i \le b \right\}.$$

Assume that the true parameter $\theta^\star$ belongs to Θ_b, every unordered pair is compared independently m times, and

$$\widehat{\theta} \in \arg\min_{\theta \in \Theta_b} L(\theta).$$

Define

$$c_b = \min_{|z| \le b} \sigma(z)(1 - \sigma(z)) > 0.$$

(a) **Optimality and Taylor expansion.** Let $\Delta = \widehat{\theta} - \theta^\star$.

(i) Prove that $\sum_i \Delta_i = 0$.

(ii) By optimality, prove $L(\theta^\star + \Delta) \le L(\theta^\star)$.

(iii) Since Θ_b is convex, every point on the line segment between $\theta^\star$ and $\widehat{\theta}$ lies in Θ_b. Apply Taylor's theorem to show that, for some point $\widetilde{\theta}$ on this segment,

$$\frac{1}{2}\Delta^\top \nabla^2 L(\widetilde{\theta})\Delta \le -\langle \nabla L(\theta^\star), \Delta\rangle.$$

(b) **A precise curvature lower bound.**

(i) Use Problem 3(d) and $\widetilde{\theta} \in \Theta_b$ to prove

$$\Delta^\top \nabla^2 L(\widetilde{\theta})\Delta \ge mc_b \sum_{i<j} (\Delta_i - \Delta_j)^2.$$

(ii) Prove the identity

$$\sum_{i<j} (\Delta_i - \Delta_j)^2 = n\|\Delta\|_2^2 - \left(\sum_i \Delta_i\right)^2.$$

(iii) Use part (a)(i) to conclude that

$$\Delta^\top \nabla^2 L(\widetilde{\theta})\Delta \ge mnc_b\|\Delta\|_2^2.$$

(c) **A precise noise upper bound.** Let $g = \nabla L(\theta^\star)$ and write $p^\star_{ij} = p_{ij}(\theta^\star)$.

(i) Show that

$$g_i = \sum_{j\neq i} (mp^\star_{ij} - Y_{ij})$$

is a sum of $n-1$ independent, mean-zero random variables. Equivalently, by writing each Y_{ij} as a sum of m Bernoulli indicators, express g_i as a sum of $m(n-1)$ bounded, mean-zero variables.

(ii) You may use the following form of Hoeffding's inequality: if $Z_1, \dots, Z_M$ are independent, mean-zero random variables with $Z_r \in [-1, 1]$, then

$$\Pr\left(\left|\sum_{r=1}^M Z_r\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{2M}\right).$$

Apply this result to one coordinate g_i, and then apply a union bound over the n coordinates, to prove that for a universal constant C, with probability at least $1-\delta$,

$$\max_i |g_i| \le C\sqrt{mn\log\frac{2n}{\delta}}.$$

You may apply a union bound over the n coordinates.

(iii) Deduce that, on the same event,

$$\|g\|_2 \le Cn\sqrt{m\log\frac{2n}{\delta}}.$$

Explain the additional factor $\sqrt{n}$ arising when the coordinatewise bound is converted to an ℓ_2 bound.

(d) **Combination.** Use Cauchy–Schwarz together with parts (a)–(c) to prove that, with probability at least $1-\delta$,

$$\boxed{\|\widehat{\theta} - \theta^\star\|_2 \le \frac{2C}{c_b}\sqrt{\frac{\log(2n/\delta)}{m}}}.$$

Numerical constants are not important; the dependence on b, n, m, and δ is.

(e) **Consequences and interpretation.** The preceding statement is precise under its stated assumptions.

(i) By what factor does the bound change when m is multiplied by four?

(ii) Show that c_b decreases as b increases. Explain quantitatively how widely separated utilities reduce likelihood curvature.

(iii) The total number of observed comparisons is $m\binom{n}{2}$. Rewrite the rate in terms of this total, ignoring logarithmic factors and constants depending on b.

(iv) For a general connected comparison graph, the complete-graph identity in part (b) is replaced by a lower bound involving the smallest nonzero eigenvalue of its graph Laplacian. Explain qualitatively why a path graph gives weaker recovery than a complete graph. This is an interpretation, not a theorem to be proved here.

(f) **Lower-bound intuition.** For two items, suppose $\Pr(1 \succ 2) = \sigma(\Delta)$.

(i) Compute $\sigma(0)$, $\sigma'(0)$, and $\sigma''(0)$. Use Taylor's theorem at zero to prove

$$\sigma(\Delta) = \frac{1}{2} + \frac{\Delta}{4} + O(\Delta^3) \qquad \text{as } \Delta \to 0.$$

(ii) Under $\Delta = 0$, one comparison is a Bernoulli observation with success probability $p_0 = 1/2$. Under $\Delta = \epsilon$, its success probability is $p_1 = \sigma(\epsilon)$. Use part (i) to prove

$$|p_1 - p_0| = \frac{|\epsilon|}{4} + O(\epsilon^3).$$

You may use without proof the following binary-testing fact: distinguishing two Bernoulli distributions whose success probabilities differ by η requires at least order $1/\eta^2$ independent observations. Substitute the expression above and conclude that distinguishing $\Delta = 0$ from $\Delta = \epsilon$ requires at least order $1/\epsilon^2$ comparisons.

(iii) Suppose an estimator could generally resolve utility gaps of size $r(m)$ after m comparisons. The preceding lower bound requires $m \gtrsim 1/r(m)^2$. Rearrange this inequality to show $r(m) \gtrsim m^{-1/2}$. Explain why this matches the dependence on m in the upper bound proved in part (d).

(g) **A numerical curvature calculation for $\mathcal{D}_0$.** The maximum possible value of $p(1-p)$ is $1/4$, attained at $p = 1/2$. Using the empirical value $\widehat{p}_{AD} = 0.9$:

(i) Compute $\widehat{p}_{AD}(1 - \widehat{p}_{AD})$ and express it as a fraction of the maximum $1/4$.

(ii) Explain the precise sense in which this pair has lower empirical curvature than a pair with win probability near $1/2$.

(iii) Distinguish the following two conclusions: the data provide strong evidence for the sign of $\theta_A - \theta_D$, but less information about the exact magnitude of this difference.