SOME INEQUALITIES WITH PROBABILITIES

DANIEL V. TAUSK

1. The inequalities

Given $p, q \in [0,1]^3$, we set:

$$P_{ij}(p,q) = p_i q_j + (1-p_i)(1-q_j), \tag{1.1}$$

for all $i, j \in \{1,2,3\}$. Our first goal is to prove the following:

1.1. **Proposition.** *For any $p, q \in [0,1]^3$, the following inequality holds:*

$$1 + \sum_{\substack{i,j=1\\ i\ne j}}^{3} P_{ij}(p,q) \ge \sum_{i=1}^{3} P_{ii}(p,q). \tag{1.2}$$

In order to prove Proposition 1.1 we need some preparatory results.

The following lemma is a particular case of a well-known result from Linear Programming, but for the reader's convenience we include here a brief proof.

1.2. **Lemma.** *If $\alpha : \mathbb{R}^n \to \mathbb{R}$ is a linear functional and $Q = \prod_{i=1}^n [a_i, b_i]$ is a rectangular block ($a_i, b_i \in \mathbb{R}$, $a_i \le b_i$, $i = 1, \ldots, n$), then the maximum of α on Q is attained at a* vertex *of Q, i.e., at an element of $\prod_{i=1}^n \{a_i, b_i\}$.*

Proof. Since α is continuous and Q is compact, we know that the maximum of α on Q is attained at some point $x \in Q$. Let:

$$I = \big\{ i \in \{1,2,\ldots,n\} : x_i = a_i \text{ or } x_i = b_i \big\}$$

and consider the subspace S of $\mathbb{R}^n$ defined by:

$$S = \big\{ u \in \mathbb{R}^n : u_i = 0, \text{ for all } i \in I \big\}.$$

If S is not contained in the kernel of α then there exists $u \in S$ with $\alpha(u) > 0$; we have $x + \varepsilon u \in Q$ for $\varepsilon > 0$ sufficiently small and $\alpha(x + \varepsilon u) > \alpha(x)$, contradicting the fact that α attains its maximum at x. Thus S is contained in the kernel of α. Let $x' \in \mathbb{R}^n$ be defined by $x'_i = x_i$, for $i \in I$ and $x'_i = a_i$, for $i \in \{1,2,\ldots,n\} \setminus I$. Then x' is a vertex of Q, $x - x' \in S$ and hence $\alpha(x') = \alpha(x)$. □

1.3. **Corollary.** *If $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form and Q is a rectangular block then B attains its maximum on $Q \times Q$ at a point (x, y) such that x and y are vertices of Q.*

Date: February 18th, 2008.

Proof. Since B is continuous and $Q \times Q$ is compact, the maximum of B on $Q \times Q$ is attained at some point $(x, y) \in Q \times Q$. By Lemma 1.2, the maximum of the linear functional $B(x, \cdot)$ on Q is attained at some vertex y' of Q; then:

$$B(x, y) \leq B(x, y').$$

Similarly, the maximum of the linear functional $B(\cdot, y')$ on Q is attained at some vertex x' of Q, so that:

$$B(x, y') \leq B(x', y').$$

But $(x', y') \in Q \times Q$ and $B(x, y) \leq B(x', y')$ imply $B(x', y') = B(x, y)$; hence the maximum of B on $Q \times Q$ is attained at the point (x', y'). □

1.4. **Corollary.** *Consider the bilinear form $B : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined by:*

$$B(x, y) = \sum_{i=1}^{3} x_i y_i - \sum_{\substack{i,j=1 \\ i \neq j}}^{3} x_i y_j,$$

for all $x, y \in \mathbb{R}^3$. Then:

$$B(x, y) \leq \tfrac{5}{4},$$

for all x, y in the cube $Q = \left[-\frac{1}{2}, \frac{1}{2}\right]^3$.

Proof. By Corollary 1.3, the maximum of B on $Q \times Q$ is attained at a point (x, y) such that x and y are vertices of Q. Since B is invariant by permutations of the coordinates and $B(x, y) = B(-x, -y)$, we can assume that $x = \frac{1}{2}(1, 1, 1)$ or $x = \frac{1}{2}(1, 1, -1)$. A direct inspection of possibilities shows that the maximum is attained at $x = y = \frac{1}{2}(1, 1, -1)$ and that $B(x, y) = \frac{5}{4}$. □

We are now ready for:

Proof of Proposition 1.1. Set $x_i = p_i - \frac{1}{2}$, $y_i = q_i - \frac{1}{2}$, $i = 1, 2, 3$, so that:

$$P_{ij}(p, q) = \left(\tfrac{1}{2} + x_i\right)\left(\tfrac{1}{2} + y_j\right) + \left(\tfrac{1}{2} - x_i\right)\left(\tfrac{1}{2} - y_j\right) = \tfrac{1}{2} + 2x_i y_j,$$

for all $i, j \in \{1, 2, 3\}$ and x, y belong to the cube $\left[-\frac{1}{2}, \frac{1}{2}\right]^3$. Then:

$$\sum_{i=1}^{3} P_{ii}(p, q) - \sum_{\substack{i,j=1 \\ i \neq j}}^{3} P_{ij}(p, q) = -\tfrac{3}{2} + 2B(x, y) \leq -\tfrac{3}{2} + 2 \cdot \tfrac{5}{4} = 1,$$

where B is defined in Corollary 1.4. The conclusion follows. □

Now denote by Q the unitary cube $[0, 1]^3$ and let μ be an arbitrary probability measure[1] on the cartesian product $Q^2 = Q \times Q$. We set:

$$\mathcal{P}_{ij}(\mu) = \int_{Q^2} P_{ij}(p, q)\, \mathrm{d}\mu(p, q), \tag{1.3}$$

for all $i, j \in \{1, 2, 3\}$.

[1] We endow $Q \times Q$ with the Borel σ-algebra.

1.5. **Proposition.** *Given a probability measure μ on Q^2, the following inequality holds:*

$$1 + \sum_{\substack{i,j=1 \\ i \neq j}}^{3} \mathcal{P}_{ij}(\mu) \geq \sum_{i=1}^{3} \mathcal{P}_{ii}(\mu).$$

Proof. Follows immediately by taking the integral with respect to μ on both sides of inequality (1.2). □

2. An experiment

Assume that we have a pair of boxes, call them A and B. Each box has three buttons, numbered 1, 2, 3 and two small lamps; one lamp is colored green and the other is colored red. If we press one of the buttons in a box then one (and only one) of the colored lamps lights up. The two boxes are manufactured by the same source (let's call it the *factory*); then, box A is sent to a far away location and box B is sent to another far away location. Each box will be received by an experimenter that will randomly choose one of the three buttons and press it. The experiment is repeated a large number of times, each time with a new pair of boxes. Each time the pair of boxes is prepared, the factory chooses a pair $p, q \in Q = [0,1]^3$. The triple of probabilities $p = (p_1, p_2, p_3)$ is programmed into the box A as follows: when button number $i \in \{1,2,3\}$ is pressed, the green lamp lights up with probability p_i and the red lamp lights up with probability $1 - p_i$. Similarly, the triple $q = (q_1, q_2, q_3)$ is programmed into the box B. The factory chooses the pairs $(p, q) \in Q^2$ randomly according to some probability measure μ on $Q^2 = Q \times Q$. It seems that this procedure is general enough so that essentially any manufacturing strategy used by the factory is covered, *as long as the boxes A and B are not allowed to interact after a button is pressed by one of the experimenters.* This "impossibility of interaction" hypothesis — let's call it *locality* — is encoded in our formalism in the hypothesis that the lottery p used by box A to associate a button to a lamp is independent of the lottery q used by box B to associate a button to a lamp. Notice that, since p and q are chosen by the same source, we allow the choice of q to be dependent on the choice of p; this is encoded in our formalism in the fact that the probability measure μ is *arbitrary*, i.e., it is not necessarily a product $\mu_1 \times \mu_2$ of two probability measures μ_1, μ_2 on Q.

Now, let $p, q \in Q$ be given and assume that the boxes A and B are programmed respectively with the triples p and q. Given, $i, j \in \{1,2,3\}$, if button number i is pressed in box A and button number j is pressed in box B, the probability that both boxes will light up the same color is $P_{ij}(p,q)$ (recall (1.1)). Now recall that the experiment is repeated a large number N of times. Given $i, j \in \{1,2,3\}$, let N_{ij} denote the number of times that button number i was pressed in box A and button number j was pressed in box B. Let N_{ij}^* denote the number of times that button number i was pressed in box A, button number j was pressed in box j and the same color lit up in both boxes. *We now assume that the experimenters choose the buttons they press using a lottery that is independent*[2] *of the lottery* μ *used by the factory to choose the pair* (p, q). This hypothesis implies that the quotient of N^*_{ij} by N_{ij} is approximately equal to $\mathcal{P}_{ij}(\mu)$ (recall (1.3)); such approximation gets better as N gets larger.

Now, assume that the experiment has been performed a large number N of times and that we have observed:

$$\frac{N^*_{ij}}{N_{ij}} \approx \frac{1}{4}, \quad \text{for } i, j \in \{1, 2, 3\},\ i \neq j, \tag{2.1}$$

$$\frac{N^*_{ii}}{N_{ii}} \approx 1, \quad \text{for } i \in \{1, 2, 3\}, \tag{2.2}$$

where $\approx$ denotes[3] "approximately equal". Is this possible? We have:

$$\mathcal{P}_{ij}(\mu) \approx \frac{1}{4},\ \text{for } i \neq j, \quad \mathcal{P}_{ii}(\mu) \approx 1,$$

so that:

$$1 + \sum_{\substack{i,j=1 \\ i \neq j}}^{3} \mathcal{P}_{ij}(\mu) \approx 1 + \frac{6}{4} = \frac{5}{2},$$

$$\sum_{i=1}^{3} \mathcal{P}_{ii}(\mu) \approx 3.$$

But this contradicts Proposition 1.5.

Now, it is a well-known fact that such an experiment can be performed[4] and that both (2.1) and (2.2) are observed. What is going on? Since these observations contradict Proposition 1.5, we have to drop one of our hypotheses: either the "non clairvoyant factory hypothesis" or the "locality hypothesis". It seems that locality is gone for good!

References

[1] A. Aspect, P. Grangier, G. Roger, Phys. Rev. Lett. **47**, 460 (1981).
[2] A. Aspect, P. Grangier, G. Roger, Phys. Rev. Lett. **49**, 91 (1982).
[3] A. Aspect, J. Dalibard, G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

---

[2]We call this the *non clairvoyant factory* hypothesis: the factory cannot predict how the buttons to be pressed (far away in space and in the future) will be chosen.

[3]How good the approximation must be? Good enough so that I can get to the contradiction with Proposition 1.5 explained below.

[4]See [1, 2, 3]. The experiments confirm the quantum predictions for measurements of polarization of entangled pairs of photons. The theoretical prediction gives $N^*_{ii} = N_{ii}$ and $\mathcal{P}_{ii}(\mu) = 1$, but we can never count on real life experiments to give us perfect correlations, so my argument allows $\mathcal{P}_{ii}(\mu) \approx 1$.