Center for Mathematical Studies in Economics and Management Science
Northwestern University, Evanston, Illinois 60208

Discussion Paper No. 1189

LARGE POISSON GAMES

by

Roger B. Myerson*

June 1997

Abstract. Existence of equilibria is proven for Poisson games with compact type sets and finite action sets. Then three theorems are introduced for characterizing limits of probabilities in Poisson games when the expected number of players becomes large. The magnitude theorem characterizes the rate at which probabilities of events go to zero. The offset theorem characterizes the ratios of probabilities of events that differ by a finite additive translation. The hyperplane theorem estimates probabilities of hyperplane events. These theorems are applied to derive formulas for pivot probabilities in binary elections, and to analyze a voting game that was studied by Ledyard.

Acknowledgments. Support from the National Science Foundation grant SES-9308139 and from the Dispute Resolution Research Center is gratefully acknowledged.

*Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2009.
E-mail: myerson@nwu.edu

LARGE POISSON GAMES

by Roger B. Myerson

1. Introduction

This paper develops some fundamental mathematical tools for analyzing games with a very large number of players, such as the game played by the voters in a large election. In such games, it is unrealistic to assume that every player knows all the other players in the game; instead, a more realistic model should admit some uncertainty about the number of players in the game. Furthermore, if we assume that such uncertainty about the number of players in the game can be described by a Poisson distribution, then the special properties of the Poisson distribution may actually make our analysis of the game simpler than under the questionable assumption that the exact number of players was common knowledge.

In a previous paper by this author (Myerson, 1994a), fundamental principles for analyzing general games with population uncertainty have been introduced, and it has shown that some convenient simplifying properties (independent actions and environmental equivalence) are uniquely satisfied by Poisson games with population uncertainty. In this paper, we will focus on some general theorems that facilitate the analysis of large Poisson games.

In Section 2, a general model of Poisson games is formulated, and existence of equilibrium is proven. Section 3 develops some general formulas that can be useful for characterizing the limits of equilibria of Poisson games as the expected number of players goes to infinity. The main results in Section 3 are the magnitude theorem which enables us to easily characterize the relative orders of magnitude of the probabilities of events, the offset theorem which characterizes the ratios of probabilities of events that differ by a finite additive translation, and the hyperplane theorem which gives probabilities of linear events. Section 4 uses these limit theorems to derive the formula for pivot probabilities in large binary elections. An application of these formulas to a voting game studied by Ledyard (1984) is developed in Section 5. The proofs of the limit theorems from Section 3 are presented in Section 6.

2. Poisson games and their equilibria

In a Poisson game, we assume that the number of players is a random variable drawn from a Poisson distribution with some mean n. (See Haight, 1967, and Johnson and Kotz, 1969.) Given this parameter n, the probability that there are k players in the game is

$$p(k|n) = e^{-n} n^k/k!$$

From the perspective of any one player in the game, the number of other players in the game (not counting this player) is also a Poisson random variable with the same mean n. This property of Poisson games is called environmental equivalence; see Myerson (1994a) for a formal derivation. To understand this environmental-equivalence property of Poisson games, imagine that you are a player in a game with population uncertainty. The number of players other than you is one less than the number of all players; but the fact that you have been recruited as a player in the game is itself evidence in favor of a larger number of players. These two effects exactly cancel out in the case where the number of players has been drawn from a Poisson distribution. That is, after learning that you are a player in a Poisson game, your posterior probability distribution on the number of other players is the same as an outside observer's prior distribution on the number of all players.

The private information of each player in the game is (or her) type, which is a random variable drawn from some given set of possible types T. In this paper, we assume that this type set T is a compact metric space. The previous paper (Myerson, 1994a) assumed a finite type set T. The class of compact metric spaces includes any finite set, as well as any closed and bounded subset of a finite-dimensional vector space; so more generality is being allowed here.

Each player's type is independently drawn from this type set T according to some given probability distribution which we denote by r. That is, for any set S that is a Borel-measurable subset of T, we let r(S) denote the probability that any given player's type is in S, and this probability is assumed to be independent of the number and types of all other players. By the decomposition property of the Poisson distribution (see Myerson, 1994a), the total number of players with types in the subset S is also a Poisson random variable with mean nr(S), and this random variable is independent of the numbers of players with types in any other disjoint sets.

Each player in the game must choose an action from a set of possible actions which we denote by C. In this paper, we assume that this action set C is a nonempty finite set.

The action profile of a group of players is the vector that lists, for each action c, the number of players in this group who are choosing action c. We let Z(C) denote the set of possible action profiles for the players in a Poisson game. That is, Z(C) is the set of vectors $x = (x(c))_{c \in C}$, with components indexed on the actions in C, such that each component x(c) is a nonnegative integer. Notice that Z(C) is a countable set, because C is finite.

The utility payoff to each player in a Poisson game depends on his type, his action, and the numbers of other players who choose each action. So utility payoffs can be mathematically specified by a utility function of the form $U:Z(C)\times C\times T \to \mathbb{R}$. Here $U(x,b,t)$ denotes the utility payoff to a player whose type is t and who chooses action b, when x is the action profile of the other players in the game (that is when, for each c in C, there are x(c) other players who choose action c, not counting this player in the case of c = b). We assume here that $U(\bullet,\bullet,\bullet)$ is a bounded function and $U(x,b,\bullet)$ is a continuous function on the type set T, for every x in Z(C) and every b in C.

These parameters (T, n, r, C, U) together define a Poisson game. For other related models of population uncertainty see also Myerson (1994a, 1994b) and Milchtaich (1997).

The strategic behavior of players in a Poisson game can be described by a distributional strategy, following Milgrom and Weber (1985). A distributional strategy for a Poisson game (T, n, r, C, U) is any probability distribution over the set $C\times T$ such that the marginal distribution on T is equal to r. So if τ is a distributional strategy then, for any action c in C and any set S that is a Borel-measurable subset of the type set T, $\tau(c,S)$ can be interpreted as the probability that a randomly sampled player will have a type in the set S and will choose the action c. Because the game specifies that players' types are drawn from the distribution r, the marginal distribution of τ on the type set T is required to satisfy the equation

$$\sum_{c\in C} \tau(c,S) = r(S)$$

for every set S that is a Borel measurable subset of T. Also, as a measure, τ must be countably additive on measurable partitions of T.

Any distributional strategy τ is associated with a unique strategy function σ that specifies numbers $\sigma(c|S)$ such that

$$\sigma(c|S) = \tau(c,S)/r(S),$$

for any measurable set of types S that has positive probability, and for any action c in C. Here $\sigma(c|S)$ can be interpreted as the conditional probability that a randomly-sampled player will choose the action c given that the player's type is in the set S. In other papers (Myerson 1994a, 1994b), strategy functions are used instead of distributional strategies to characterize players' behavior in a Poisson game, but it will be more convenient here to use distributional strategies.

We may let $\Delta(C)$ denote the set of probability distributions on the finite action set C. Any distributional strategy τ induces a marginal probability distribution on C, which may also be denoted by τ without danger of confusion. That is, under any distribution strategy τ, the marginal probability $\tau(c)$ of any action c in C is

$$\tau(c) = \tau(c,T).$$

When the players behave according to the distributional strategy τ (or the corresponding strategy function σ), the number of players who choose each action c in C is a Poisson random variable with mean $n\tau(c)$. Furthermore, the number of players who choose the action c is independent of the numbers of players who choose all other actions. This result is called the <u>independent-actions</u> property, and it can be shown to characterize Poisson games (see Myerson, 1994a). So for any x in Z(C), the probability that x is the action profile of the players in the game is

$$P(x|n\tau) = \prod_{c \in C} \left(\frac{e^{-n\tau(c)} \, (n\tau(c))^{x(c)}}{x(c)!} \right).$$

By the environmental-equivalence property of Poisson games, any player in the game assesses the same probabilities for the action profile of the other players in the game (not counting himself). Thus, the expected payoff to a player of type t who chooses action b, when the other players are expected to behave according to the distributional strategy τ, is

$$\sum_{x \in Z(C)} P(x|n\tau)\ U(x,b,t)$$

Let $G(b,n\tau)$ denote the set of all types for whom choosing action b would maximize this expected payoff over all possible actions, when n is the expected number of players and τ is the distributional strategy. That is,

$$G(b,n\tau) = \left\{t \in T \mid \sum_{x \in Z(C)} P(x|n\tau)U(x,b,t) = \max_{c \in C} \sum_{x \in Z(C)} P(x|n\tau)U(x,c,t)\right\}.$$

This set $G(b,n\tau)$ is a closed subset of T, because it is defined by an equality among two continuous functions of t.

A distributional strategy τ is an <u>equilibrium</u> of the Poisson game iff

$$\tau(b,G(b,n\tau)) = \tau(b),\ \ \forall b \in C.$$

That is, a distributional strategy is an equilibrium iff, for every action b, all the probability of choosing action b comes from types for whom b is an optimal action, when everyone else is expected to behave according to this distributional strategy.

Our first main result is a general existence theorem for equilibria of Poisson games. (The existence theorem of Myerson 1994a allows forms of population uncertainty more general than the Poisson, but only allows finite type sets, whereas infinite type sets are allowed here.)

<u>Theorem 0.</u> For any Poisson game (T, n, r, C, U) as above (where T is a compact metric space, C is a finite set, and U is continuous and bounded), there must exist at least one distributional strategy that is an equilibrium.

<u>Proof</u>. We use a fixed-point argument on $\Delta(C)$, the set of probability distributions on the

finite action set C. Notice that the definitions of $P(x|n\tau)$ and $G(b,n\tau)$ above depended only on the components $(\tau(c))_{c\in C}$, which form a vector in $\Delta(C)$. So $G(b,n\eta)$ is well defined for any vector $\eta = (\eta(c))_{c\in C}$ in $\Delta(C)$.

For any vector η in $\Delta(C)$, let $R^*(\eta)$ denote the set of distributional strategies τ that satisfy the equation $\tau(c) = \tau(c,G(b,n\eta))$ for every c in C. Let $R(\eta)$ denote the set of all vectors $(\tau(c))_{c\in C}$ in $\Delta(C)$ such that τ is a distributional strategy in $R^*(\eta)$, where we use the convention $\tau(c) = \tau(c,T)$. These sets $R^*(\eta)$ and $R(\eta)$ are convex, because they are defined by linear conditions on τ.

The sets $R^*(\eta)$ and $R(\eta)$ are also nonempty. To show this, put an arbitrary ordering on the finite set C and consider the distributional strategy τ such that

$$\tau(c,S) = r(\{t\in S|\ c = \min\{b|\ t\in G(b,n\eta)\}\}).$$

This distributional strategy τ assigns all type-t players to the minimal action (according to our ordering) among their optimal responses to the anticipated behavior η. Then this distributional strategy τ is in $R^*(\eta)$, and the vector $(\tau(c))_{c\in C}$ is in $R(\eta)$.

$G(b,n\eta)$ is a closed subset of T and depends upper-hemicontinuously on η, because the probabilities $P(x|n\eta)$ are continuous functions of η and the utility numbers $U(x,c,t)$ are bounded. Now suppose that we are given sequences $\{\eta_k\}_{k=1}^{\infty}$ and $\{\tau_k\}_{k=1}^{\infty}$ such that $\tau_k \in R^*(\eta_k)$ for every k, and suppose that $\eta_k \to \eta$ as $k \to \infty$. The set of distributional strategies on the compact set $C\times T$ is itself a compact metric space (see Milgrom and Weber, 1985, and Billingsley, 1968), and so there must exist an infinite subsequence in which τ_k converges to some distributional strategy τ in the weak topology on measures. For every action b we have $\tau_k(b,T\backslash G(b,n\eta_k)) = 0$ for each k, and so $\tau(b,T\backslash G(b,n\eta)) = 0$. The limit vector $(\tau(c))_{c\in C}$ must therefore be in $R(\eta)$, and so

$R:\Delta(C)\rightarrow\rightarrow\Delta(C)$ is an upper-hemicontinuous correspondence.

Thus, by the Kakutani fixed-point theorem, there exists some η in $\Delta(C)$ such that $\eta \in R(\eta)$. The distributional strategy in $R^*(\eta)$ that verifies this inclusion is an equilibrium. Q.E.D.

3. Limits of probabilities in large Poisson games

We now develop some general theorems for estimating probabilities of events in equilibria of large Poisson games. Let us consider a sequence of Poisson games that are parameterized by the expected size parameter n. For each of these games, suppose that some equilibrium τ_n has been identified that predicts what the players' behavior would be in the game. Our goal is to characterize the limits of probabilities in these equilibria as the size parameter n goes to infinity. In this section, we will not actually use the full Poisson-game structure (T, n, r, C, U) that was introduced in the previous section. We will only use the set of actions C and the size parameter n, along with the corresponding equilibrium τ_n.

For each n and each c in C, $\tau_n(c)$ is defined such that $n\tau_n(c)$ is the expected number of players who would choose action c in the predicted equilibrium of the game of size n. The size parameter n denotes the expected total number of players in the game; that is,

$$\sum_{c \in C} n\tau_n(c) = n, \text{ and so } \sum_{c \in C} \tau_n(c) = 1, \ \forall n.$$

In this section, we can let τ_n denote the vector $(\tau_n(c))_{c \in C}$ in $\Delta(C)$, because the other components of the distributional strategy that was denoted by τ in the preceding section will not be used here. The vector $n\tau_n = (n\tau_n(c))_{c \in C}$ may be called the expected results vector in the game of size n.

To approximate Poisson probabilities, we may use Stirling's formula (see Abramowitz and Stegun, 1965), one version of which asserts that n! is approximately equal to $(n/e)^n\sqrt{2\pi n + \pi/3}$ when n is large. To be more precise, let us define

$$\iota(k) = \frac{k!}{(k/e)^k\sqrt{2\pi k + \pi/3}},$$

for any nonnegative integer k. Then we have

$$\lim_{k\to\infty} \iota(k) = 1.$$

This convergence is quite rapid and actually satisfies the stronger condition $\lim_{k\to\infty} \iota(k)^k = 1$. Even for low values of k, $\iota(k)$ is not far from 1: $\iota(0) = .977$, $\iota(1) = 1.004$, $\iota(2) = 1.001$ and the difference of $\iota(k)$ from 1 becomes less than 0.0001 when k=9.

Thus, with expected results vector $n\tau_n$, the Poisson probability of any possible action profile x in Z(C) is

$$P(x|n\tau_n) = \prod_{c\in C}\left(\frac{e^{-n\tau_n(c)}\,(n\tau_n(c))^{x(c)}}{x(c)!}\right)$$

$$= \prod_{c\in C}\left(\frac{e^{-n\tau_n(c)}\,(n\tau_n(c))^{x(c)}}{\iota(x(c))\,(x(c)/e)^{x(c)}\,\sqrt{2\pi x(c) + \pi/3}}\right)$$

$$= \prod_{c\in C}\left(\frac{e^{x(c) - x(c)\log(x(c)/(n\tau_n(c))) - n\tau_n(c)}}{\iota(x(c))\,\sqrt{2\pi x(c) + \pi/3}}\right)$$

(Here the log function is the logarithm base e.) To simplify this probability formula, let us define the function $\psi:\mathbb{R}_+ \to \mathbb{R}$ by the equations

$$\psi(\theta) = \theta(1 - \log(\theta)) - 1 = -\int_1^\theta \log(\gamma)\,d\gamma, \ \forall\theta > 0,$$

$$\psi(0) = \lim_{\theta\to 0} \psi(\theta) = -1.$$

It is straightforward to verify that

$$\psi(1) = 0, \;\; \psi(\theta) < 0, \;\; \forall\theta \neq 1,$$

and $\psi(\bullet)$ is a concave function with derivative

$$\psi'(\theta) = -\log(\theta).$$

The graph of this ψ function is shown in Figure 1.

[Insert Figure 1 about here.]

With this ψ function, the Poisson probabilities can be written

$$\text{(3.1)} \qquad P(x|n\tau_n) = \prod_{c\in C} \left(\frac{e^{n\tau_n(c)\,\psi(x(c)/(n\tau_n(c)))}}{\iota(x(c))\,\sqrt{2\pi x(c) + \pi/3}} \right).$$

To make equation (3.1) valid in the case where $\tau_n(c)=0$, we adopt the following convention:

$$\text{if } \tau_n(c) = 0 \text{ and } x(c) = 0 \text{ then } \tau_n(c)\,\psi(x(c)/(n\tau_n(c))) = 0,$$

$$\text{if } \tau_n(c) = 0 \text{ and } x(c) > 0 \text{ then } \tau_n(c)\,\psi(x(c)/(n\tau_n(c))) = -\infty$$

(and $e^{-\infty} = 0$). Taking the logarithm of equation (3.1), we get

$$\text{(3.2)} \qquad \log(P(x_n|n\tau_n))/n$$

$$= \sum_{c\in C} \tau_n(c)\,\psi(x_n(c)/(n\tau_n(c))) + \frac{\sum_{c\in C} \left(\log(\iota(x_n(c))) + .5\log(2\pi x_n(c) + \pi/3)\right)}{n}.$$

Let us say that a sequence of vectors $\{x_n\}_{n=1}^{\infty}$ in Z(C) has a <u>magnitude</u> μ iff the sequence $\log(P(x_n|n\tau_n))/n$ converges to μ as n goes to infinity. That is, μ is the magnitude of the sequence $\{x_n\}_{n=1}^{\infty}$ iff

$$\mu = \lim_{n\to\infty} \log(P(x_n|n\tau_n))/n.$$

Notice that this magnitude must be zero or negative, because the logarithm of a probability is never positive. When the magnitude μ is negative, the probabilities $P(x_n|n\tau_n)$ are going to zero at the rate of $e^{\mu n}$. The following lemma, which follows easily from equation (3.2), is useful for computing magnitudes of sequences. This lemma and all the other theorems and corollaries of this section are proven in Section 6.

Lemma 1. Let $\{x_n\}_{n=1}^{\infty}$ be any sequence of possible action profiles in Z(C). Then

$$\lim_{n\to\infty} \frac{\log(P(x_n|n\tau_n))}{n} = \lim_{n\to\infty} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c)}{n\tau_n(c)}\right).$$

That is, if either of these two limits exists, then both limits exist and are equal.

We now extend the notion of magnitude to sequences of events. We can represent any event A as a subset of Z(C). The probability of the event A when $n\tau_n$ is the expected results vector is

$$P(A|n\tau_n) = \sum_{x\in A} P(x|n\tau_n).$$

Given any sequence of events $\{A_n\}_{n=1}^{\infty}$ such that $A_n \subseteq Z(C)$ for each n, we may say that the magnitude of the sequence $\{A_n\}_{n=1}^{\infty}$ is

$$\lim_{n\to\infty} \log(P(A_n|n\tau_n))/n,$$

whenever this limit exists.

Let us say that $\{x_n\}_{n=1}^{\infty}$ is a major sequence of points in the event-sequence $\{A_n\}_{n=1}^{\infty}$ iff each x_n is a point in A_n and the sequence of points $\{x_n\}_{n=1}^{\infty}$ has a magnitude that is equal to the greatest magnitude of any sequence that can be selected from the A_n events; that is

$$x_n \in A_n \ \ \forall n, \text{ and}$$

$$\lim_{n\to\infty} \log(P(x_n|n\tau_n))/n = \lim_{n\to\infty} \max_{y\in A_n} \log(P(y|n\tau_n))/n.$$

To satisfy the definition of a major sequence, we require that the limits in the above equation must exist. The following theorem asserts that the magnitude of any sequence of events must coincide with the magnitude of any major sequence of points in these events.

Theorem 1. A sequence of events $\{A_n\}_{n=1}^{\infty}$ has a magnitude if and only if there exists a major sequence of points in $\{A_n\}_{n=1}^{\infty}$; and if such a major sequence exists then the magnitude of $\{A_n\}_{n=1}^{\infty}$ is equal to the magnitude of any major sequence in $\{A_n\}_{n=1}^{\infty}$. That is,

$$\lim_{n\to\infty} \log(P(A_n|n\tau_n))/n = \lim_{n\to\infty} \max_{y_n\in A_n} \log(P(y_n|n\tau_n))/n.$$

Theorem 1 and Lemma 1 imply as a corollary that, in large Poisson games, almost all of the probability in any event must be concentrated in the regions where the formula

$$\sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x(c)}{n\,\tau_n(c)}\right)$$

is close to its maximum. If $B_n \subseteq A_n$ for all n then the hypothesis about $\{B_n\}_{n=1}^{\infty}$ in the following corollary is equivalent to assuming that every major sequence of points in $\{A_n\}_{n=1}^{\infty}$ can have only finitely many points that are in the corresponding subsets $\{B_n\}_{n=1}^{\infty}$.

Corollary 1. Suppose that $\{A_n\}_{n=1}^{\infty}$ is a sequence of events that has a finite magnitude. Suppose that $\{B_n\}_{n=1}^{\infty}$ is a sequence of events such that

$$\limsup_{n\to\infty} \max_{y_n\in B_n} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{y_n(c)}{n\,\tau_n(c)}\right) < \lim_{n\to\infty} \max_{x_n\in A_n} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c)}{n\,\tau_n(c)}\right).$$

Then $\lim_{n\to\infty} \dfrac{P(B_n|n\tau_n)}{P(A_n|n\tau_n)} = 0$ and $\lim_{n\to\infty} \dfrac{P(A_n\backslash B_n|n\tau_n)}{P(A_n|n\tau_n)} = 1.$

Lemma 1 and Corollary 1 alert us to a useful way of recalibrating action profiles. For any possible action profile x in Z(C), for any action c in C, the ratio $x(c)/(n\tau_n(c))$ may be called the c-offset corresponding to x when $n\tau_n$ is the expected results vector. That is, the c-offset is a ratio which describes the number of players who are choosing c as a fraction of the mean of the Poisson distribution from which this number was drawn.

For any action c in C, we may say that $\alpha(c)$ is the limit of major c-offsets in the sequence of events $\{A_n\}_{n=1}^{\infty}$ iff, for every major sequence of points $\{x_n\}_{n=1}^{\infty}$ in $\{A_n\}_{n=1}^{\infty}$, we have

$$\alpha(c) = \lim_{n\to\infty} x_n(c)/(n\tau_n(c)),$$

and (to avoid triviality) there exists at least one such major sequence of points in $\{A_n\}_{n=1}^{\infty}$.

Consider any vector $w = (w(c))_{c\in C}$ in $\mathbb{R}^C$ such that each component w(c) is an integer (which may be positive or negative or zero). For any event A, we let A−w denote the set of vectors in Z(C) such that adding the vector w would yield a vector in the event A; that is,

$$A-w = \{x-w \mid x \in A,\ x-w \in Z(C)\}.$$

The following theorem relates the probabilities of such pairs of events that differ by such an additive translation in large Poisson games, when limits of major offsets exist.

Theorem 2. Let w be any vector in $\mathbb{R}^C$ such that each component w(c) is an integer. For each action c such that $w(c) \neq 0$, suppose that $\lim_{n\to\infty} n\tau_n(c) = +\infty$, and suppose that some number $\alpha(c)$ is the limit of major c-offsets in the sequence of events $\{A_n\}_{n=1}^{\infty}$. Then

$$\lim_{n\to\infty} \frac{P(A_n - w \mid n\tau_n)}{P(A_n \mid n\tau_n)} = \prod_{c\in C} \alpha(c)^{w(c)}.$$

Our definitions of magnitude and major sequence can be applied to a single event A as well as to a sequence of events $\{A_n\}_{n=1}^{\infty}$ in the obvious way. That is, given $A \subseteq Z(C)$, μ is the magnitude of the A and $\{x_n\}_{n=1}^{\infty}$ is a major sequence in A iff $x_n \in A$ for all n and

$$\mu = \lim_{n\to\infty} \log(P(A|n\tau_n))/n = \lim_{n\to\infty} \log(P(x_n|n\tau_n))/n.$$

Theorem 1 may be called the magnitude theorem, and Theorem 2 may be called the offset theorem. If the magnitude of an event A is larger than the magnitude of some other event B, then we know that the probability of B will become infinitesimal relative to the probability of A, and the conditional probability of B given $A \cup B$ will go to 0 as $n \to \infty$. But the magnitude theorem is not useful for comparing the probabilities of the events that differ by adding or subtracting a fixed vector, because the difference between such events may seem small in large Poisson games and so they usually have the same magnitude. So relative probabilities of events that differ by a simple additive translation must be compared using the offset theorem instead.

The magnitude of an event only tells us about the rate at which its probability goes to zero. To estimate the probability of any individual point more precisely, we can apply equation (3.1), using the fact that the ι factor is close to 1 and therefore can be ignored. Our next limit theorem gives estimates of the probabilities of other events, but it requires some further restrictions. This limit theorm only considers a fixed event that has a simple linear structure, and the distributional strategy τ_n is assumed to be a constant τ that is independent of n.

Let J be a positive integer. Let $w_1,\ldots,w_J$ be vectors such that, for each i, $w_i = (w_i(c))_{c\in C}$ is a vector in $\mathbb{R}^C$ and each component $w_i(c)$ is an integer. We allow that $w_i(c)$ may be a negative integer (in which case w_i would not be in Z(C), because Z(C) only includes the nonnegative integer vectors in $\mathbb{R}^C$). Suppose that the vectors $w_1,\ldots,w_J$ are linearly independent, in the sense that the vector equation $\gamma_1 w_1 + \ldots + \gamma_J w_J = \vec{0}$ has no solutions other than $\gamma_1 = \ldots = \gamma_J = 0$.

Let $H(w_1,\ldots,w_J)$ denote the set of all vectors x in Z(C) such that there exist integers

$\gamma_1,...,\gamma_J$ such that $x = \gamma_1 w_1 + ... + \gamma_J w_J$. (Notice that the integers $\gamma_1,...,\gamma_J$ may be negative, but the linear combination $\gamma_1 w_1 + ... + \gamma_J w_J$ must have all nonnegative components to be Z(C).) This set $H(w_1,...,w_J)$ may be called the <u>hyperplane event</u> in Z(C) that is spanned by the basis $\{w_1,...,w_J\}$. That is,

$$H(w_1,...,w_J) = \left\{\sum_{i=1}^{J} \gamma_i w_i \,\middle|\, \sum_{i=1}^{J} \gamma_i w_i(c) \geq 0 \;\; \forall c, \;\; \gamma_i \text{ is an integer } \forall i\right\} \subseteq Z(C).$$

Let $H^*(w_1,...,w_J)$ denote the set that we get if we drop the restriction that each γ_i coefficient must be an integer; that is

$$H^*(w_1,...,w_J) = \left\{\sum_{i=1}^{J} \gamma_i w_i \,\middle|\, \sum_{i=1}^{J} \gamma_i w_i(c) \geq 0 \;\; \forall c, \;\; \gamma_i \in \mathbb{R} \;\; \forall i\right\} \subseteq \mathbb{R}^C.$$

The following theorem tells us how to estimate the probabilities of hyperplane events in large Poisson games.

<u>Theorem 3.</u> Suppose that $\tau_n = \tau$ for every n, and $\tau(c) > 0$ for all c in C. Let y be the vector that maximizes $\sum_{c \in C} \tau(c)\psi(y(c)/\tau(c))$ subject to the constraint that $y \in H^*(w_1,...,w_J)$. Let M be the J×J matrix such that the (i,j) entry is

$$M_{ij} = \sum_{c \in C} w_i(c) w_j(c)/y(c).$$

Then

$$\lim_{n \to \infty} \frac{P(H(w_1,..w_J) \mid n\tau)}{n^{(J-\#C)/2}\, e^{\sum_{c \in C} n\tau(c)\psi(y(c)/\tau(c))}} = \sqrt{\frac{(2\pi)^J}{\det(M) \prod_{c \in C}(2\pi y(c))}}.$$

To read this theorem more clearly, let us use the approximate equality symbol ≈ to indicate functions of n whose ratio converges to 1 as n goes to infinity. With this notation, the conclusion of Theorem 3 may be rewritten

$$P(H(w_1,..w_J) \mid n\tau) \approx \frac{(2\pi n)^{J/2}\, e^{\sum_{c \in C} n\tau(c)\psi(y(c)/\tau(c))}}{\sqrt{\det(M)} \prod_{c \in C}\sqrt{2\pi n y(c)}}.$$

Let y_n denote the point in $H(w_1,...,w_J)$ that is closest to the vector ny described in Theorem 3. Using equation (3.1) and the first-order conditions of the optimization that defines y, the above approximate inequality can be simplified to

$$P(H(w_1,...,w_J)|n\tau) \approx \left(\frac{(2\pi n)^{J/2}}{\sqrt{\det(M)}}\right) P(y_n|n\tau) \ .$$

In the proof (see Section 6), Theorem 3 is actually derived from this approximate equality.

4. Pivot probabilities in voting games

As an application of the limit theorems from the preceding section, consider a model of a large election in which the voters must choose among two candidates who are numbered 1 and 2. The winner will be the candidate with the most votes. In the case of a tie, the winner will be determined by a fair coin toss.

Consider a sequence of election games, parameterized by the size n which will go to infinity in the limit. In the game of size n, suppose that our equilibrium analysis has predicted that the numbers of votes for candidates 1 and 2 will be independent Poisson random variables with means $n\tau_n(1)$ and $n\tau_n(2)$ respectively. The environmental-equivalence property of Poisson games implies that each voter similarly perceives that the number of other voters who vote for candidates 1 and 2 are independent Poisson random variables with means $n\tau_n(1)$ and $n\tau_n(2)$ respectively. Let us assume here that these $\tau_n(c)$ numbers converge to some limit that we denote by $\tau(c)$; that is,

$$\lim_{n\to\infty} \tau_n(c) = \tau(c), \ \ \forall c \in C.$$

(If these probabilities did not converge then a convergent subsequence could be chosen, so there is little loss of generality in assuming such convergence.) When we consider only the two actions of voting for 1 and voting for 2, we could take C = {1,2} and have $\tau_n(1) + \tau_n(2) = 1$ for all n. However, to derive formulas that can be readily applied when voters have more alternatives (such as abstention or voting for other candidates), we use here only the weaker assumption that at least one limiting $\tau(c)$ is positive; that is,

$$\tau(1) + \tau(2) > 0.$$

In the analysis of rational voting behavior, we may want to estimate a voter's probability of changing the outcome of the election, which is called his pivot probability. So let us consider a voter who is planning to vote for candidate 1. There are two ways that his vote for candidate 1 could change the outcome of the election. His vote could break a tie in which candidate 1 would have lost the fair coin toss, or his vote could make a tie in which candidate 1 would win the fair coin toss. Thus, the pivot probability of a vote for candidate 1 is

$$.5 \times \text{P(Tie)} + .5 \times \text{P(Candidate 1 is one vote behind candidate 2)} .$$

To compute this pivot probability, we begin by computing the probability of a tie. In set-theoretic terms, the event of a tie is $\{x \in Z(C) | \; x(1)=x(2)\}$, which is the hyperplane event spanned by the single vector ω such that

$$\omega(1) = \omega(2) = 1$$

In the notation of the preceding section, with this vector ω, the event of a tie is denoted $H(\omega)$.

To assess the magnitude of a tie, Theorem 1 and Lemma 1 tell us to characterize the major sequences $\{x_n\}_{n=1}^{\infty}$ which asymptotically maximize

$$\tau_n(1)\,\psi\!\left(\frac{x_n(1)}{n\tau_n(1)}\right) + \tau_n(2)\,\psi\!\left(\frac{x_n(2)}{n\tau_n(2)}\right)$$

subject to the constraint $x_n \in H(\omega)$, that is

$$x_n(2) = x_n(1).$$

Choosing a subsequence if necessary, suppose that the x_n sequence gives us limiting offsets $\alpha(c) = \lim_{n\to\infty} x_n(c)/(n\tau_n(c))$ for each c. Then the above maximand converges to

$$\tau(1)\psi(\alpha(1)) + \tau(2)\psi(\alpha(2)),$$

and the constraint on x_n becomes

$$\tau(1)\alpha(1) = \tau(2)\alpha(2).$$

Because the derivative $\psi'(\alpha(c))$ equals $-\log(\alpha(c))$, the maximum over α subject to this constraint is achieved when

$$0 = -\tau(1)\log(\alpha(1)) - \tau(2)\log\big(\tau(1)\alpha(1)/\tau(2)\big)\big(\tau(1)/\tau(2)\big),$$

and this first-order condition is satisfied together with the constraint at the unique solution

$$\alpha(1) = \sqrt{\tau(2)/\tau(1)},\quad \alpha(2) = \sqrt{\tau(1)/\tau(2)}.$$

Thus, $\{x_n\}_{n=1}^{\infty}$ is a major sequence of points in the event of a tie iff

$$\lim_{n\to\infty} x_n(1)/\big(n\tau_n(1)\big) = \alpha(1) = \sqrt{\tau(2)/\tau(1)},$$

$$\lim_{n\to\infty} x_n(2)/\big(n\tau_n(2)\big) = \alpha(2) = \sqrt{\tau(1)/\tau(2)}.$$

So for each c, this number $\alpha(c)$ is the limit of major c-offsets in the event of a tie.

In any such major sequence, the magnitude (from Lemma 1) is

$$\lim_{n\to\infty} \tau_n(1)\,\psi\left(\frac{x_n(1)}{n\tau_n(1)}\right) + \tau_n(2)\,\psi\left(\frac{x_n(2)}{n\tau_n(2)}\right)$$

$$= \tau(1)\,\psi\left(\sqrt{\tau(2)/\tau(1)}\right) + \tau(2)\,\psi\left(\sqrt{\tau(1)/\tau(2)}\right)$$

$$= \tau(1)\left(\sqrt{\tau(2)/\tau(1)}\,(1 - \log(\sqrt{\tau(2)/\tau(1)})) - 1\right)$$

$$+ \tau(2)\left(\sqrt{\tau(1)/\tau(2)}\,(1 - \log(\sqrt{\tau(1)/\tau(2)})) - 1\right)$$

$$= 2\sqrt{\tau(1)\,\tau(1)} - \tau(1) - \tau(2) = -(\sqrt{\tau(1)} - \sqrt{\tau(2)})^2.$$

Thus, by the magnitude theorem (Theorem 1),

$$\text{(4.1)} \qquad \lim_{n\to\infty} \log(P(H(\omega)\,|\,n\tau_n))/n = 2\sqrt{\tau(1)\,\tau(2)} - \tau(1) - \tau(2).$$

For each candidate c, let $v(c\,|\,n\tau_n)$ denote the pivot probability of a vote for candidate c. We have seen that the pivot probability of a vote for candidate 1 is half of the probability of a tie plus half the probability of candidate 1 being behind by one vote. But the event of candidate 1 being behind by one vote differs from the event of a tie by subtracting one vote for candidate 1. So by the offset theorem, the probability of candidate 1 being behind by one vote is approximately $\sqrt{\tau(2)/\tau(1)}$ multiplied by the probability of a tie. Thus, the pivot probability of a vote for candidate 1 is

$$\text{(4.2)} \qquad v(1\,|\,n\tau_n) \approx P(H(\omega)\,|\,n\tau_n)\,\left(1 + \sqrt{\tau(2)/\tau(1)}\right)/2 = P(H(\omega)\,|\,n\tau_n)\left(\frac{\sqrt{\tau(1)} + \sqrt{\tau(2)}}{2\sqrt{\tau(1)}}\right).$$

(Recall that the approximate equality symbol $\approx$ here indicates functions of n whose ratio converges to 1 as n goes to infinity.) Similarly, the pivot probability of a vote for candidate 2 is half of the probability of a tie plus half the probability of candidate 2 being behind by one vote, which is approximately

$$\text{(4.3)} \qquad v(2\,|\,n\tau_n) \approx P(H(\omega)\,|\,n\tau_n)\,\left(1 + \sqrt{\tau(1)/\tau(2)}\right)/2 = P(H(\omega)\,|\,n\tau_n)\left(\frac{\sqrt{\tau(1)} + \sqrt{\tau(2)}}{2\sqrt{\tau(2)}}\right).$$

Equations (4.2) and (4.3) imply that

$$(4.4) \qquad \lim_{n\to\infty} \frac{v(1|n\tau_n)}{v(2|n\tau_n)} = \sqrt{\frac{\tau(2)}{\tau(1)}}.$$

In particular, if the expected vote total for candidate 1 is less than the expected vote total for candidate 2, then a vote for candidate 1 is more likely to be pivotal than a vote for candidate 2, because the probability of candidate 1 being behind by one vote is greater than the probability of candidate 2 being behind by one vote.

We can apply Theorem 3 to get a stronger approximation of these pivot probabilities, but this hyperplane theorem requires that the distributional strategies τ_n must be constant. So suppose now that

$$\tau_n = \tau, \ \forall n.$$

Applying Theorem 3 to $H(\omega)$ where $\omega = (1,1)$, we have $J = 1$, and

$$H^*(\omega) = \{x \in \mathbb{R}^C | \ x(1) = x(2) \geq 0\}.$$

Theorem 3 tells us to find the vector y that maximizes $\sum_{c\in C} \tau(c)\psi(y(c)/\tau(c))$ subject to $y \in H^*(\omega)$, that is $y(1) = y(2) \geq 0$. Again using $\psi' = -\log$, we find that this maximum is achieved when

$$y(1) = y(2) \ \text{ and } \ 0 = -\log(y(1)/\tau(1)) - \log(y(1)/\tau(2)),$$

that is, when

$$y(1) = y(2) = \sqrt{\tau(1)\,\tau(2)}.$$

Thus we get $y(c) = \alpha(c)\tau(c)$ for each c, where $\alpha(c)$ is as above, and so

$$\sum_{c\in C} \tau(c)\ \psi(y(c)/\tau(c)) = \tau(1)\,\psi\left(\sqrt{\tau(2)/\tau(1)}\right) + \tau(2)\,\psi\left(\sqrt{\tau(1)/\tau(2)}\right)$$

$$= 2\sqrt{\tau(1)\,\tau(1)} - \tau(1) - \tau(2).$$

With J=1, then the matrix M is just the single number.

$$M = 1/y(1) + 1/y(2) = 2/\sqrt{\tau(1)\tau(2)}.$$

Substituting these values into the formula in Theorem 3 and simplifying, we get the following approximate formula for the probability of a tie:

$$P(H(\omega)|n\tau) \approx \frac{e^{n\left(2\sqrt{\tau(1)\tau(2)} - \tau(1) - \tau(2)\right)}}{2\sqrt{\pi n \sqrt{\tau(1)\tau(2)}}}.$$

Thus, the pivot probabilities can be approximated by the formulas

(4.5) $$v(1|n\tau) \approx \left(\frac{e^{n\left(2\sqrt{\tau(1)\tau(2)} - \tau(1) - \tau(2)\right)}}{4\sqrt{\pi n \sqrt{\tau(1)\tau(2)}}}\right)\left(\frac{\sqrt{\tau(1)} + \sqrt{\tau(2)}}{\sqrt{\tau(1)}}\right),$$

(4.6) $$v(2|n\tau) \approx \left(\frac{e^{n\left(2\sqrt{\tau(1)\tau(2)} - \tau(1) - \tau(2)\right)}}{4\sqrt{\pi n \sqrt{\tau(1)\tau(2)}}}\right)\left(\frac{\sqrt{\tau(1)} + \sqrt{\tau(2)}}{\sqrt{\tau(2)}}\right).$$

These pivot-probability formulas can also be derived from mathematical formulas involving Bessel functions. When the number of votes for candidates 1 and 2 are independent Poisson random variables with means $n\tau(1)$ and $n\tau(2)$ respectively, the probability that candidate 1 gets exactly k more votes than candidate 2 is

$$e^{-n(\tau(1)+\tau(2))}\left(\frac{\tau(1)}{\tau(2)}\right)^{k/2} I_k\left(2n\sqrt{\tau(1)\tau(2)}\right),$$

where I_k is a modified Bessel function (see formula 9.6.10 in Abramowitz and Stegun, 1965). With large n, the following approximation formula for modified Bessel functions can be applied

$$I_k\left(2n\sqrt{\tau(1)\tau(2)}\right) \approx \frac{e^{2n\sqrt{\tau(1)\tau(2)}}}{\sqrt{4\pi n\sqrt{\tau(1)\tau(2)}}}$$

(see formula 9.7.1 in Abramowitz and Stegun, 1965).

5. Ledyard's model with costly voting

To illustrate the power of these results, we now derive a Poisson version of a basic theorem in social choice that was originally shown (for a multinomial model) by Ledyard (1984). By using a Poisson model, we should be able to derive Ledyard's results more cleanly and simply than was possible with the model of nonrandom population size that Ledyard (1984) used.

Consider a voting game in which the players are voters who can choose to abstain from voting. In this voting game, each player's type has two components: his policy type and his voting cost. Suppose that the set of possible policy types is some finite set Θ, and suppose that the voting costs are drawn out of the interval from 0 to 1. So the type set T is the compact set

$$T = \Theta \times [0,1].$$

As in Section 4, there are two candidates numbered 1 and 2, but we now allow that each player has three possible actions denoted by elements in the set $C = \{0,1,2\}$. Here action 1 is voting for candidate 1, action 2 is voting for candidate 2, and action 0 is abstaining. As above, the winner is the candidate with the most votes, and we assume that the winner will be determined by the toss of a fair coin in the event of a tie.

Each player's policy type θ in Θ determines the policy benefits $u(c,\theta)$ that he will get if candidate c is the winner of the election. But we must also take the cost of voting into account. When candidate c wins, a player who has policy type θ and voting cost γ would get a total utility payoff equal to $u(c,\theta)-\gamma$ if he voted in the election, while a similar player would get a total utility payoff equal to $u(c,\theta)$ if he abstained in the election.

Let the number of players in this voting game be a Poisson random variable with mean n. Each player's policy type is a random variable drawn from Θ according to some probability distribution ρ, where $\rho(\theta)$ denotes the probability of having policy type θ. Each player's voting cost is a random variable drawn from [0,1] according to a probability distribution that has a cumulative distribution function F such that the derivative at zero F'(0) is strictly positive. That is, we assume that the probability density of voting costs must be strictly positive at 0, but nobody can have a negative cost of voting.

We also assume that the policy types and voting costs of all players are independent random variables. That is, each player's policy type and voting cost are independent of each other and of all other players' types.

The total utility payoffs defined above are bounded and depend continuously on the voter's type, as the equilibrium-existence theorem in Section 2 requires. Thus, this Poisson game of size n has at least one equilibrium, which we may denote by τ_n. The main result of this section is that, if the expected number of players in the voting game is large, then the candidate who offers the greater expected policy benefits will almost surely win in equilibrium.

<u>Theorem 4.</u> In the voting game described above, suppose that the expected policy benefits for a randomly-sampled voter are greater from candidate 2 than from candidate 1; that is,

$$\sum_{\theta\in\Theta} \rho(\theta)\, u(2,\theta) > \sum_{\theta\in\Theta} \rho(\theta)\, u(1,\theta).$$

Then the probability of candidate 2 winning in the voting game of size n under the equilibrium τ_n must converge to 1 as the size parameter n goes to infinity.

<u>Proof</u>. Let Θ_1 denote the set of policy types in Θ that prefer candidate 1, and let Θ_2 denote the other policy types; that is

$$\Theta_1 = \{\theta\in\Theta \mid u(1,\theta) > u(2,\theta)\}, \quad \Theta_2 = \{\theta\in\Theta \mid u(1,\theta) \le u(2,\theta)\}.$$

In equilibrium, for each candidate c, each player with policy type in Θ_c will either vote for his preferred candidate c or abstain, because voting for the less preferred candidate is strictly dominated by abstaining.

Given any candidate c in C, let $-c$ denote the other candidate in C. Let $v_n(c)$ denote the probability that a vote for candidate c would be pivotal in the equilibrium τ_n of the voting game of size n. In this equilibrium, a player of policy type θ in Θ_c prefers to actually vote for candidate c (rather than abstain) iff his voting cost is less than

$$\big(u(c,\theta) - u(-c,\theta)\big)\, v_n(c).$$

Thus, the probability that a randomly sampled player will vote for candidate c in equilibrium, which we denote by $\tau_n(c)$, must satisfy the equation

$$\tau_n(c) = \sum_{\theta\in\Theta_c} \rho(\theta) F\Big(\big(u(c,\theta)-u(-c,\theta)\big)v_n(c)\Big). \tag{5.1}$$

We now claim that the expected total number of votes $n\tau_n(1)+n\tau_n(2)$ must go to infinity as $n\to\infty$. If not, then both candidates' expected scores $n\tau_n(1)$ and $n\tau_n(2)$ would have finite limits (taking a subsequence if necessary), and then the pivot probabilities $v_n(1)$ and $v_n(2)$ would converge to the positive pivot probabilities that are associated with independent Poisson-distributed vote totals that have these limiting expected values. But then (5.1) would imply that $\tau_n(2)$ must have a strictly positive limit, and so $n\tau_n(2)$ goes to infinity, as claimed.

So if we look only at the players who actually vote, then the sequence of games considered here has an expected voting population that goes to infinity as $n\to\infty$. That is, we could reparameterize our sequence by letting

$$m(n) = n\tau_n(1)+n\tau_n(2) \text{ and } \hat{\tau}_{m(n)}(c) = \tau_n(c)/m(n), \forall c\in\{1,2\},$$

and we would then get a sequence $\{\hat{\tau}_m\}$ that satisfies all the properties of the $\{\tau_n\}$ sequence in Section 4 (where we required that a positive fraction of the population must be voting for 1 and 2). So equation (4.4) from Section 4 can be applied here to give us

$$(5.2) \qquad \lim_{n\to\infty} \frac{v_n(1)}{v_n(2)} = \lim_{n\to\infty} \sqrt{\frac{\tau_n(2)}{\tau_n(1)}}.$$

The fact that at least one candidate's expected score is going to infinity implies that both pivot probabilities $v_n(1)$ and $v_n(2)$ go to 0 as $n\to\infty$. Thus, by differentiability of the cumulative distribution function F at zero,

$$\sum_{\theta\in\Theta_c} \rho(\theta) F\big((u(c,\theta)-u(-c,\theta))v_n(c)\big) \approx \sum_{\theta\in\Theta_c} \rho(\theta) F'(0)\big(u(c,\theta)-u(-c,\theta)\big)v_n(c))$$

for each candidate c. Then (5.1) gives us

$$\frac{\tau_n(1)}{\tau_n(2)} \approx \frac{\sum_{\theta\in\Theta_1} \rho(\theta) F'(0)(u(1,\theta)-u(2,\theta))v_n(1)}{\sum_{\theta\in\Theta_2} \rho(\theta) F'(0)(u(2,\theta)-u(1,\theta))v_n(2)}.$$

So applying (5.2) we get

$$\frac{\tau_n(1)}{\tau_n(2)} \approx \frac{\sum_{\theta\in\Theta_1} \rho(\theta)(u(1,\theta)-u(2,\theta))}{\sum_{\theta\in\Theta_2} \rho(\theta)(u(2,\theta)-u(1,\theta))} \sqrt{\frac{\tau_n(2)}{\tau_n(1)}}$$

That is,

$$(5.3) \qquad \frac{\tau_n(1)}{\tau_n(2)} \approx \left(\frac{\sum_{\theta\in\Theta_1} \rho(\theta)(u(1,\theta)-u(2,\theta))}{\sum_{\theta\in\Theta_2} \rho(\theta)(u(2,\theta)-u(1,\theta))}\right)^{2/3}.$$

By the basic assumption that candidate 2 offers greater expected policy benefits than candidate 1, the right-hand side of (5.3) is strictly less than one. So the expected score of candidate 2 must be going to infinity and, in the limit, the expected score of candidate 1 is less than the expected score of candidate 2 by a strictly positive fraction of candidate 2's expected score.

Recall that the standard deviation of any Poisson random variable is the square root of its expected value, and this square root is a vanishing fraction of the expected value as the expected value becomes large. So the expected excess of candidate 2's score over candidate 1's score is becoming infinitely many times the standard deviation of either score as $n \to \infty$. Thus, the probability that candidate 2 wins must be converging to one. Q.E.D.

Following Ledyard (1984), we can now take the story back one stage to the point in time where the candidates choose their policy positions. Suppose that the players in the voting game have preferences over some given policy space, and each candidate can choose any policy position in this space. After the candidates choose these policies, the policy benefits $u(c,\theta)$ will be equal to the benefits that a player of policy type θ would get from the policy position chosen by candidate c. Theorem 4 tells us that, when n is large, any candidate who does not choose a policy position that maximizes the players' expected benefits can be beaten almost surely by a candidate who chooses a policy position that maximizes the players' expected benefits. Thus, both candidates should rationally choose a policy position that maximizes the players' expected benefits. If there is a unique policy position that maximizes the players' expected benefits, then both candidates must rationally choose that same position, in which case nobody will actually vote in the voting game. Thus Ledyard (1984) showed that democracy may achieve the classical utilitarian ideal of expected welfare maximization in a voting game where nobody actually votes in equilibrium!

6. Proofs of the limit theorems

We begin with a useful fact about $\psi(\theta) = \theta(1-\log(\theta))-1$. For any nonnegative number θ,

(6.1) $\qquad \psi(\theta) < 2 - \theta.$

To verify this inequality, it can be shown by differentiation that the convex function $2-\theta-\psi(\theta)$ is minimal when $\theta = e$, where is equal to $3-e$, which is positive.

Lemma 1. Let $\{x_n\}_{n=1}^{\infty}$ be any sequence of possible action profiles in Z(C). Then

$$\lim_{n\to\infty} \frac{\log(P(x_n|n\tau_n))}{n} = \lim_{n\to\infty} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c)}{n\tau_n(c)}\right).$$

That is, if either of these two limits exists, then both limits exist and are equal.

Proof. From equation (3.1) we get

$$\log(P(x_n|n\tau_n))/n - \sum_{c\in C} \tau_n(c)\,\psi\big(x_n(c)/(n\tau_n(c))\big)$$
$$= \sum_{c\in C} \big(\log(\iota(x_n(c))) + .5\log(2\pi x_n(c) + \pi/3)\big)/n$$

The term $\log(\iota(x(c)))/n$ is must go to zero as n goes to infinity, because $\iota(x(c))$ is always close to 1, and the term $\log(2\pi x_n(c) + \pi/3)$ is always positive. So the equality in Lemma 1 can fail only if there exists some action c, some positive number ε, and some infinite subsequence of the n's such that

$$\log(2\pi x_n(c) + \pi/3) > \varepsilon n \ \text{ and } \ x_n(c) > e^{\varepsilon n}/(2\pi) - 1/6\,,\quad \forall n,$$

and so $x_n(c)/n$ goes to $+\infty$. Inequality (6.1) implies

$$\tau_n(c)\psi\big(x_n(c)/(n\tau_n(c))\big) < 2\tau_n(c) - x_n(c)/n,$$

and $$\tau_n(c)\psi(x_n(c)/(n\tau_n(c))) + .5\log(2\pi x_n(c) + \pi/3)/n$$

$$< 2\tau_n(c) - (x_n(c) - .5\log(2\pi x_n(c) + \pi/3))/n.$$

With $x_n(c)/n$ going to $+\infty$ and $\tau_n(c)$ bounded, the right-hand sides of these two inequalities must both go to $-\infty$ as $n \to +\infty$. But then $\sum_{c\in C} \tau_n(c)\,\psi(x_n(c)/(n\tau_n(c)))$ and $\log(P(x_n|n\tau_n))/n$ must both go to $-\infty$. That is, in any subsequence where the difference between $\sum_{c\in C} \tau_n(c)\,\psi(x_n(c)/(n\tau_n(c)))$ and $\log(P(x_n|n\tau_n))/n$ goes to any limit other than zero, both expressions must go to $-\infty$. Thus their limits must be equal. Q.E.D.

For any integer k and any nonnegative number λ, let $p(k|\lambda)$ denote the probability that a Poisson random variable with mean λ would equal k. That is,

$$p(k|\lambda) = e^{-\lambda}\,\lambda^k/k!$$

We now prove two more computational lemmas about Poisson distributions.

Lemma 2. Let λ be any nonnegative number and let i be any integer such that $i > \lambda$. Consider the event that a Poisson random variable with mean λ is greater than or equal to i. The probability of this event can be bounded by the following inequality

$$\sum_{k=i}^{\infty} p(k|\lambda) \le p(i|\lambda)\left(\frac{i}{i-\lambda}\right)$$

Proof. For any positive integer δ,

$$p(i+\delta|\lambda) = e^{-\lambda}\,\lambda^{i+\delta}/(i+\delta)! = (e^{-\lambda}\,\lambda^i/i!)\,(\lambda^\delta/((i+1)\ldots(i+\delta))) \le p(i|\lambda)\,(\lambda/i)^\delta.$$

Thus,

$$\sum_{k=i}^{\infty} p(k|\lambda) \le p(i|\lambda)\left(\sum_{\delta=0}^{\infty}(\lambda/i)^{\delta}\right) = p(i|\lambda)\left(\frac{1}{1-\lambda/i}\right) = p(i|\lambda)\left(\frac{i}{i-\lambda}\right).$$

Q.E.D.

Lemma 3. Let λ be any positive number, and let h and k be any two integers. then there exists some numbers ε and η such that ε is between 0 and 1, η is between h+ε and k+ε, and

$$\frac{p(k|\lambda)}{p(h|\lambda)} = e^{\lambda(\psi((k+\varepsilon)/\lambda)-\psi((h+\varepsilon)/\lambda))} = \left(\frac{\lambda}{h+\varepsilon}\right)^{k-h} e^{-(k-h)^2/(2\eta)}.$$

Proof. Suppose first that k < h. Then

$$\log\left(\frac{p(k|\lambda)}{p(h|\lambda)}\right) = \log\left(\frac{\lambda^k\ h!}{k!\ \lambda^h}\right) = \sum_{i=k+1}^{h} \log(i/\lambda).$$

But from the basic definition of the Riemann integral,

$$\lambda\int_{k/\lambda}^{h/\lambda} \log(\theta)\,d\theta \le \sum_{i=k+1}^{h} \log(i/\lambda) \le \lambda\int_{(k+1)/\lambda}^{(h+1)/\lambda} \log(\theta)\,d\theta,$$

because the log function is monotone increasing. So there must exist some ε between 0 and 1 such that

$$\log\left(\frac{p(k|\lambda)}{p(h|\lambda)}\right) = \lambda\int_{(k+\varepsilon)/\lambda}^{(h+\varepsilon)/\lambda} \log(\theta)\,d\theta = \lambda\left(\psi((k+\varepsilon)/\lambda) - \psi((h+\varepsilon)/\lambda)\right),$$

where the second equality follows from the fact that $\psi'(\theta) = -\log(\theta)$. Reversing the roles of k and h, it is straightforward to show that these equalities also hold in the case where k > h.

By a second-order Taylor expansion, there is some number η between k and h such that

$$\psi\left(\frac{k+\varepsilon}{\lambda}\right) - \psi\left(\frac{h+\varepsilon}{\lambda}\right) = -\log\left(\frac{h+\varepsilon}{\lambda}\right)\left(\frac{k-h}{\lambda}\right) - \frac{1}{2}\left(\frac{\lambda}{\eta}\right)\left(\frac{k-h}{\lambda}\right)^2.$$

The second equality in the Lemma follows easily from this Taylor expansion. Q.E.D.

Theorem 1. A sequence of events $\{A_n\}_{n=1}^{\infty}$ has a magnitude if and only if there exists a major sequence of points in $\{A_n\}_{n=1}^{\infty}$; and if such a major sequence exists then the magnitude of $\{A_n\}_{n=1}^{\infty}$ must be equal to the magnitude of any major sequence in $\{A_n\}_{n=1}^{\infty}$. That is,

$$\lim_{n\to\infty} \log(P(A_n|n\tau_n))/n = \lim_{n\to\infty} \max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n.$$

Proof. For any negative number ν, define the set $S(\nu,n)$ such that

$$S(\nu,n) = \left\{x\in Z(C) \middle| \sum_{c\in C} \tau_n(c)\,\psi\big(x(c)/(n\tau_n(c))\big) < \nu\right\}.$$

Our first claim is that $\limsup_{n\to\infty} \log(P(S(\nu,n)|n\tau_n))/n$ is not greater than ν.

To prove this claim, we cover $S(\nu,n)$ by #C+1 subsets. (Here #C denotes the number of actions in the set C.) For any action c in C, let $S_c(\nu,n)$ be the set such that

$$S_c(\nu,n) = \{x\in S(\nu,n)|\ x(c) \geq n(2-\nu)\}.$$

Also, let $S^*(\nu,n)$ be the set such that

$$S^*(\nu,n) = \{x\in S(\nu,n)|\ x(c) < n(2-\nu),\ \forall c\in C\}.$$

Thus, $S(\nu,n) \subseteq S^*(\nu,n) \cup (\cup_{c\in C} S_c(\nu,n))$.

Let θ denote the next integer larger than $n(2-\nu)$. By (6.1), the inequality $\theta \geq n(2-\nu)$ implies that

$$\tau_n(c)\ \psi(\theta/(n\tau_n(c))) < \tau_n(c)(2-\theta/(n\tau_n(c))) \leq 2\tau_n(c) - (2-\nu) \leq \nu.$$

So the probability that exactly θ players choose c satisfies

$$p(\theta|n\tau_n(c)) = \frac{e^{n\tau_n(c)\psi(\theta/(n\tau_n(c)))}}{\iota(\theta)\sqrt{2\pi\theta+\pi/3}} < e^{n\nu}.$$

Our set $S_c(\nu,n)$ is a subset of the event that at least θ players choose action c. So by Lemma 2,

$$P(S_c(\nu,n)|n\tau_n) < e^{n\nu}\left(\frac{n(2-\nu)}{n(2-\nu)-n\tau_n(c)}\right) < e^{n\nu}\left(\frac{2-\nu}{1-\nu}\right).$$

The set $S^*(\nu,n)$ contains at most $(n(2-\nu))^{\#C}$ points, each of which has a probability less than $e^{n\nu}$. So

$$P(S^*(\nu,n)|n\tau_n) < e^{n\nu}\,(n(2-\nu))^{\#C}.$$

Thus we get

$$(6.2) \qquad P(S(\nu,n)|n\tau_n) < \left(n(2-\nu))^{\#C} + \#C\,(2-\nu)/(1-\nu)\right)e^{n\nu},$$

which proves our first claim.

Now suppose that $\{A_n\}_{n=1}^{\infty}$ is a sequence of events that has at least one major sequence of points, and let μ denote the magnitude of such a major sequence in $\{A_n\}_{n=1}^{\infty}$; that is,

$$\mu = \lim_{n\to\infty} \max_{y\in A_n} \log(P(y|n\tau_n))/n.$$

Let ε be any strictly positive number. Then for all sufficiently large n, the set A_n must be a subset of $S(\mu+\varepsilon,n)$. So for all sufficiently large n, we get

$$P(A_n|n\tau_n) < \left(n(2-\mu-\varepsilon))^{\#C} + \#C\,(2-\mu-\varepsilon)/(1-\mu-\varepsilon)\right)e^{n(\mu+\varepsilon)} < e^{n(\mu+2\varepsilon)},$$

because $e^{n\varepsilon} > n(2-\mu-\varepsilon))^{\#C} + \#C\,(2-\mu-\varepsilon)/(1-\mu-\varepsilon)$ when n is large. Thus,

$$\limsup_{n\to\infty} \log(P(A_n|n\tau_n)/n \le \mu + 2\varepsilon,\ \forall\varepsilon > 0,$$

which in turn implies that

$$\limsup_{n\to\infty} \log(P(A_n|n\tau_n)/n \le \mu.$$

But the assumption that a major sequence $\{x_n\}_{n=1}^{\infty}$ in $\{A_n\}_{n=1}^{\infty}$ has magnitude μ also implies that

$$\liminf_{n\to\infty} \log(P(A_n|n\tau_n))/n \ge \lim_{n\to\infty} \log(P(x_n|n\tau_n))/n = \mu,$$

because each x_n is in A_n. Thus we conclude

$$\lim_{n\to\infty} \log(P(A|n\tau_n))/n = \mu.$$

That is, if there is a major sequence in $\{A_n\}_{n=1}^{\infty}$, then its magnitude is equal to the magnitude of $\{A_n\}_{n=1}^{\infty}$.

Now suppose that the sequence of events $\{A_n\}_{n=1}^{\infty}$ has a magnitude μ. Obviously, no subsequence of points in $\{A_n\}_{n=1}^{\infty}$ can have a magnitude greater than μ, and so

$$\mu \geq \limsup_{n\to\infty} \max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n .$$

If the $\liminf_{n\to\infty} \max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n$ were strictly less than μ, then we could choose an infinite subsequence in which the numbers $\max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n$ converge to this limit-infimum. But along this subsequence, a major sequence of points would exist, and so (as just shown) the limit of $\max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n$ would be equal to the limit of $\log(P(A_n|n\tau_n)/n$, which equals μ. Thus, $\lim_{n\to\infty} \max_{y_n \in A_n} \log(P(y_n|n\tau_n))/n$ must exist and must equal μ. Q.E.D.

<u>Corollary 1.</u> Suppose that $\{A_n\}_{n=1}^{\infty}$ is a sequence of events that has a finite magnitude. Suppose that $\{B_n\}_{n=1}^{\infty}$ is a sequence of events such that

$$\limsup_{n\to\infty} \max_{y_n \in B_n} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{y_n(c)}{n\,\tau_n(c)}\right) < \lim_{n\to\infty} \max_{x_n \in A_n} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c)}{n\,\tau_n(c)}\right).$$

Then $\lim_{n\to\infty} \dfrac{P(B_n|n\tau_n)}{P(A_n|n\tau_n)} = 0$ and $\lim_{n\to\infty} \dfrac{P(A_n\backslash B_n|n\tau_n)}{P(A_n|n\tau_n)} = 1$.

<u>Proof</u>. Let μ denote the magnitude of $\{A_n\}_{n=1}^{\infty}$. If the corollary failed then we could find some infinite subsequence along which $P(B_n|n\tau_n)/P(A_n|n\tau_n)$ is bounded below by some positive number q, and so

$$0 \geq \liminf_{n\to\infty} \log(P(B_n|n\tau_n))/n \geq \lim_{n\to\infty} \left(\log(q) + \log(P(A_n|n\tau_n))\right)/n$$

$$= \lim_{n\to\infty} \log(P(A_n|n\tau_n))/n = \mu.$$

So this subsequence could also be chosen so that the $\{B_n\}$ subsequence has a magnitude and

$$\lim_{n\to\infty} \log(P(B_n|n\tau_n))/n \geq \mu.$$

By Theorem 1, we could then select points y_n in B_n such that, along this subsequence,

$$\lim_{n\to\infty} \log(P(y_n|n\tau_n))/n \geq \mu,$$

and then Lemma 1 would imply

$$\lim_{n\to\infty} \sum_{c\in C} \tau(c)\psi\left(y_n(c)/(n\tau_n(c))\right) \geq \mu.$$

But this result would contradict the strict inequality that was assumed in the corollary. Q.E.D.

Theorem 2. Let w be any vector in $\mathbb{R}^C$ such that each component w(c) is an integer. For each action c such that $w(c) \neq 0$, suppose that $\lim_{n\to\infty} n\tau_n(c) = +\infty$, and suppose that some number $\alpha(c)$ is the limit of major c-offsets in the sequence of events $\{A_n\}_{n=1}^{\infty}$. Then

$$\lim_{n\to\infty} \frac{P(A_n - w|n\tau_n)}{P(A_n|n\tau_n)} = \prod_{c\in C} \alpha(c)^{w(c)}.$$

Proof. Let ε be any positive number. Let $D_n(\varepsilon)$ be the set of all x in A_n such that

$$\alpha(c) - \varepsilon < \frac{x(c) - |w(c)|}{n\tau_n(c)} \quad \text{and} \quad \frac{x(c) + |w(c)|}{n\tau_n(c)} < \alpha(c) + \varepsilon$$

for every c such that $w(c) \neq 0$. Because $w(c)/(n\tau_n(c))$ converges to 0 and $\alpha(c)$ is the limit of major c-offsets in $\{A_n\}_{n=1}^{\infty}$ for each such c, any major sequence in $\{A_n\}_{n=1}^{\infty}$ must have at most finitely many points outside of $D_n(\varepsilon)$. So by Corollary 1,

$$\lim_{n\to\infty} \frac{P(D_n(\varepsilon)|n\tau_n)}{P(A_n|n\tau_n)} = 1.$$

Let μ denote the magnitude of $\{A_n\}_{n=1}^{\infty}$. Because all major sequences in $\{A_n\}_{n=1}^{\infty}$ are eventually in $\{D_n(\varepsilon)\}_{n=1}^{\infty}$, we know that μ is also the magnitude of $\{D_n(\varepsilon)\}_{n=1}^{\infty}$. If $\{x_n - w\}_{n=1}^{\infty}$ is any sequence of points in $\{(A_n - w) \setminus (D_n(\varepsilon) - w)\}_{n=1}^{\infty}$ then

$$\limsup_{n\to\infty} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c) - w(c)}{n\tau_n(c)}\right) = \limsup_{n\to\infty} \sum_{c\in C} \tau_n(c)\ \psi\left(\frac{x_n(c)}{n\tau_n(c)}\right) < \mu,$$

because ψ is continuous and $n\tau_n(c) \to +\infty$ whenever $w(c) \neq 0$. So by Corollary 1,

$$\lim_{n\to\infty} \frac{P((A_n - w)\setminus(D_n(\varepsilon) - w)|n\tau_n)}{P(D_n(\varepsilon)|n\tau_n)} = 0.$$

Thus,

$$\lim_{n\to\infty} \frac{P(A_n - w|n\tau_n)}{P(A_n|n\tau_n)} = \lim_{n\to\infty} \frac{P(A_n - w|n\tau_n)}{P(D_n(\varepsilon)|n\tau_n)} = \lim_{n\to\infty} \frac{P(D_n(\varepsilon) - w|n\tau_n)}{P(D_n(\varepsilon)|n\tau_n)}.$$

Now consider any point $x - w$ in $D_n(\varepsilon) - w$ and the corresponding point x in $D_n(\varepsilon)$. The ratio of the probabilities of these two points is

$$\frac{P(x-w|n\tau_n)}{P(x|n\tau_n)} = \prod_{c\in C} \frac{e^{-n\tau_n(c)}(n\tau_n(c))^{x(c)-w(c)}/(x(c)-w(c))!}{e^{-n\tau_n(c)}(n\tau_n(c))^{x(c)}/x(c)!}$$

$$= \prod_{c\in C} \left(\frac{x(c)!/(x(c)-w(c))!}{(n\tau_n(c))^{w(c)}}\right).$$

If $w(c) > 0$ then $x(c)!/(x(c)-w(c))!$ is the product of $w(c)$ factors between $x(c)$ and $x(c)-w(c)$. Similarly, if $w(c) < 0$ then $(x(c)-w(c))!/x(c)!$ is the product of $-w(c)$ factors between $x(c)$ and $x(c)-w(c)$. Applying the definition of $D_n(\varepsilon)$, we then get

$$\prod_{c\in C} (\alpha(c)-\varepsilon)^{w(c)} \leq \frac{P(x-w|n\tau_n)}{P(x|n\tau_n)} \leq \prod_{c\in C} (\alpha(c)+\varepsilon)^{w(c)}.$$

Now there are two cases to consider. First, consider the case where, for all sufficiently large n, for every point x in $D_n(\varepsilon)$, the point $x-w$ has all nonnegative components and so is in $Z(C)$. Then for all sufficiently large n, $P(D_n(\varepsilon)-w|n\tau_n)$ and $P(D_n(\varepsilon)|n\tau_n)$ are sums of point probabilities that can be put in a one-to-one correspondence where each corresponding pair has a ratio between $\prod_{c\in C}(\alpha(c)-\varepsilon)^{w(c)}$ and $\prod_{c\in C}(\alpha(c)+\varepsilon)^{w(c)}$. So we get

$$\prod_{c\in C}(\alpha(c)-\varepsilon)^{w(c)} \le \frac{P(D_n(\varepsilon)-w|n\tau_n)}{P(D_n(\varepsilon)|n\tau_n)} \le \prod_{c\in C}(\alpha(c)+\varepsilon)^{w(c)}.$$

Because these inequalities hold for all ε, we can conclude

$$\lim_{n\to\infty} \frac{P(A_n-w|n\tau_n)}{P(A_n|n\tau_n)} = \prod_{c\in C}\alpha(c)^{w(c)}$$

which proves the theorem for this case.

Now consider the alternative case where there exist arbitrarily large n such that, for some point x in $D_n(\varepsilon)$, the point $x-w$ has some negative components and so is not in $Z(C)$. In this case, the argument in the preceding paragraph fails only because there may be some extra terms in $P(D_n(\varepsilon)|n\tau_n)$ that do not correspond to any terms in $P(D_n(\varepsilon)-w|n\tau_n)$. Thus, we can only claim

$$\frac{P(D_n(\varepsilon)-w|n\tau_n)}{P(D_n(\varepsilon)|n\tau_n)} \le \prod_{c\in C}(\alpha(c)+\varepsilon)^{w(c)}.$$

Because this condition holds for any positive ε, we get

$$0 \le \lim_{n\to\infty} \frac{P(A_n-w|n\tau_n)}{P(A_n|n\tau_n)} \le \prod_{c\in C}\alpha(c)^{w(c)}.$$

But notice that this case can occur only if there is some action c such that $w(c) > 0$ and $\alpha(c) = 0$, because otherwise the condition $x \in D_n(\varepsilon)$ would force $x(c)$ to be larger than $w(c)$ for all sufficiently large n. So in this case we can also conclude that

$$\lim_{n\to\infty} \frac{P(A_n-w|n\tau_n)}{P(A_n|n\tau_n)} = \prod_{c\in C}\alpha(c)^{w(c)} = 0. \qquad \text{Q.E.D.}$$

Theorem 3. Suppose that $\tau_n = \tau$ for every n, and $\tau(c) > 0$ for all c in C. Let y be the vector that maximizes $\sum_{c\in C} \tau(c)\psi(y(c)/\tau(c))$ subject to the constraint that $y \in H^*(w_1,...,w_J)$. Let M be the J×J matrix such that the (i,j) entry is

$$M_{ij} = \sum_{c\in C} w_i(c)w_j(c)/y(c).$$

Then

$$\lim_{n\to\infty} \frac{P(H(w_1,..w_J)\,|\,n\,\tau)}{n^{(J-\#C)/2}\, e^{\sum_{c\in C} n\,\tau(c)\psi(y(c)/\tau(c))}} = \sqrt{\frac{(2\,\pi)^J}{\det(M)\,\prod_{c\in C}(2\,\pi\,y(c))}}.$$

Proof. First we verify that the vector y in the theorem is well defined. The feasible set $H^*(w_1,...,w_J)$ is a closed convex subset of $\mathbb{R}^C$. The objective $\sum_{c\in C} \tau(c)\psi(y(c)/\tau(c))$ is strictly concave in y, goes to $-\infty$ if any y(c) goes to $+\infty$, and has partial derivatives $\partial/\partial y(c)$ that go to $+\infty$ when any y(c) goes to 0. Thus, there is a unique vector y that maximizes this objective within this feasible set, and the components of this vector y are all strictly positive. At this point y, the first-order conditions of the optimization problem give us

$$0 = \sum_{c\in C} w_i(c)\ \log(y(c)/\tau(c)),\ \ \forall i\in\{1,..,J\},$$

or equivalently

$$\prod_{c\in C} \left(\frac{y(c)}{\tau(c)}\right)^{w_i(c)} = 1. \tag{6.3}$$

The vector y can be written in the form $\sum_{i=1}^{J} \gamma_i w_i$, where each γ_i is a real number. For each n, let y_n be a vector in $H(w_1,...,w_J)$ that is as close as possible to ny. Such a vector y_n can be written in the form $\sum_{i=1}^{J} \hat{\gamma}_{i,n} w_i$ where each $\hat{\gamma}_{i,n}$ is an integer that differs from $n\gamma_i$ by less than 1. Thus, the quantities $|y_n(c)-ny(c)|+1$ are bounded, uniformly over all n and c, by some upper bound which we may denote by D.

These points $\{y_n\}_{n=1}^{\infty}$ form a major sequence of points in $H(w_1,...,w_J)$, because they approximate the vectors ny which maximize $\sum_{c\in C} \tau(c)\psi(x(c)/(n\tau(c)))$ over all x in the larger set $H^*(w_1,...,w_J)$. Because y is the unique solution to this strictly concave maximization, the ratio $y(c)/\tau(c)$ is the limit of major c-offsets, for each c in C. That is, a sequence $\{x_n\}_{n=1}^{\infty}$ in $H(w_1,...,w_J)$ will have strictly smaller magnitude unless, for each c, the ratios $x_n(c)/(ny(c))$ and $x_n(c)/y_n(c)$ converge to 1 as $n\to\infty$.

For any small positive number θ, let

$$\Lambda_n(\theta) = \{x\in H(w_1,...,w_J) \mid (1-\theta)\max\{y_n(c),ny(c)\} < x(c) < (1+\theta)\min\{y_n(c),ny(c)\} - 1,\ \forall c\in C\}.$$

and let $B_n(\theta) = H(w_1,...,w_J)\backslash\Lambda_n(\theta)$. Then by Corollary 1,

$$\lim_{n\to\infty} \frac{P(\Lambda_n(\theta)\mid n\tau)}{P(H(w_1,...,w_J)\mid n\tau)} = 1$$

for any strictly positive θ.

Thus, when we estimate $P(H(w_1,...,w_J)\mid n\tau)$ for large n, we only need to consider points x in $\Lambda_n(\theta)$, where θ is positive but arbitrarily close to 0. So let x be any such point in $\Lambda_n(\theta)$. Lemma 3 gives us

$$\frac{P(x\mid n\tau)}{P(y_n\mid n\tau)} = \prod_{c\in C}\left(\left(\frac{n\tau(c)}{y_n(c)+\varepsilon(c)}\right)^{x(c)-y_n(c)} e^{-(x(c)-y_n(c))^2/(2\eta(c))}\right)$$

for some numbers $\varepsilon(c)$ between 0 and 1, and $\eta(c)$ between $y_n(c)+\varepsilon(c)$ and $x(c)+\varepsilon(c)$. Notice

$$\prod_{c\in C}\left(\frac{n\tau(c)}{y_n(c)+\varepsilon_n(c)}\right)^{x(c)-y_n(c)} = \prod_{c\in C}\left(\frac{n\tau(c)}{ny(c)}\right)^{x(c)-y_n(c)}\left(\frac{ny(c)}{y_n(c)+\varepsilon(c)}\right)^{x(c)-y_n(c)}.$$

But

$$\prod_{c\in C}\left(\frac{n\tau(c)}{ny(c)}\right)^{x(c)-y_n(c)} = 1$$

by equation (6.3), because the vector $x-y_n$ is a linear combination of the w_i vectors. Applying the upper bound D for the quantities $|y_n(c)-ny(c)|+\varepsilon(c)$ and using the inequality

$\theta y_n(c) \geq |x(c)-y_n(c)|$, we also get

$$1 - \theta D \leq \prod_{c\in C} \left(\frac{ny(c)}{y_n(c)+\varepsilon(c)} \right)^{x(c)-y_n(c)} \leq \frac{1}{1-\theta D}.$$

Having $x(c)+\varepsilon(c)$ and $y_n(c)+\varepsilon(c)$ both between $(1-\theta)ny(c)$ and $(1+\theta)ny(c)$, we also get

$$e^{-(x(c)-y_n(c))^2/(2(1-\theta)ny(c))} \leq e^{-(x(c)-y_n(c))^2/(2\eta(c))} \leq e^{-(x(c)-y_n(c))^2/(2(1+\theta)ny(c))}.$$

Thus, by choosing θ arbitrarily close to 0 and n large, we can guarantee that

$P(\Lambda_n(\theta)|n\tau)/P(y_n|n\tau)$ is arbitrarily close to

(6.4) $$\sum_{x\in\Lambda_n(\theta)} e^{-\sum_{c\in C} (x(c)-y_n(c))^2/(2ny(c))}.$$

Any point x in $\Lambda_n(\theta) \subseteq H(w_1,...,w_J)$ can be written in the form

$$x(c) = y_n + \sum_{i=1}^{J} k_i w_i, \text{ where } k_1,...,k_J \text{ are all integers.}$$

When we make this substitution, formula (6.4) becomes

$$\sum_{k_1} \cdots \sum_{k_J} e^{-\sum_{c\in C} (\sum_{i=1}^{J} k_i w_i(c))^2/(2ny(c))} = \sum_{k_1} \cdots \sum_{k_J} e^{-\sum_{i=1}^{J} \sum_{j=1}^{J} M_{ij}k_ik_j/(2n)}$$

where $M_{ij} = \sum_{c\in C} w_i(c)w_j(c)/y(c)$.

Substituting $h_i = k_i/\sqrt{n}$ into the above formulas and dividing by $n^{J/2}$ we get

$$\frac{P(H(w_1,...,w_J)|n\tau)}{n^{J/2}\ P(y_n|n\tau)} \approx \frac{P(\Lambda_n(\theta)|n\tau)}{n^{J/2}\ P(y_n|n\tau)} \approx \sum_{h_1} \cdots \sum_{h_J} e^{-\sum_{i=1}^{J} \sum_{j=1}^{J} M_{ij}h_ih_j/2} \left(\frac{1}{\sqrt{n}}\right)^J$$

where the summation is over vectors $(h_1,...,h_J)$ where the components are integer multiples of $1/\sqrt{n}$ such that $y_n + \sum_i \sqrt{n}\,h_i w_i$ is in $\Lambda_n(\theta)$. (Recall that the approximate-equality symbol "$\approx$" is used here to indicate quantities whose ratio goes to 1 as $n\to\infty$.) But for any vector $(h_1,...,h_J)$ whose components are integer multiples of $1/\sqrt{n}$, the vector $y_n + \sum_i \sqrt{n}\,h_i w_i$ will be in $\Lambda_n(\theta)$ for all sufficiently large n, because

$$\lim_{n\to\infty} \frac{\sum_i \sqrt{n}\, h_i w_i(c)}{n y(c)} = 0.$$

So the definition of Riemann integration gives us

$$\lim_{n\to\infty} \frac{P(H(w_1,\ldots,w_J)\,|\,n\tau)}{n^{J/2}\ P(y_n\,|\,n\tau)} \approx \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} e^{-\sum_{i=1}^{J} \sum_{j=1}^{J} M_{ij} h_i h_j/2}\ dh_1 \ldots dh_J .$$

Multiplying the above integrand by $\sqrt{\det(M)}/(2\pi)^{J/2}$ yields a multivariate normal probability density with mean $\vec{0}$ and covariance matrix M^{-1}, and the integral of such a density is 1. Thus,

$$\lim_{n\to\infty} \frac{P(H(w_1,\ldots,w_J)\,|\,n\tau)}{n^{J/2}\ P(y_n\,|\,n\tau)} = \frac{(2\pi)^{J/2}}{\sqrt{\det(M)}} .$$

To complete the proof of the theorem, we can use equation (3.1), together with the first-order conditions of the optimization that defines y, to show that

$$1 = \lim_{n\to\infty} \frac{P(y_n\,|\,n\tau)\ \prod_{c\in C} \sqrt{2\pi y_n(c)}}{e^{\sum_{c\in C} n\tau(c)\psi(y_n(c)/(n\tau(c)))}} = \lim_{n\to\infty} \frac{P(y_n\,|\,n\tau)\ \prod_{c\in C} \sqrt{2\pi n y(c)}}{e^{\sum_{c\in C} n\tau(c)\psi(y(c)/\tau(c))}} .$$

Q.E.D.

REFERENCES

M. Abramowitz and I. Stegun, Handbook of Mathematical Tables, New York: Dover (1965).

P. Billingsley, Convergence of Probability Measures, New York: Wiley (1968).

F. A. Haight, Handbook of the Poisson Distribution, New York: Wiley (1967).

N. L. Johnson and S. Kotz, Discrete Distributions, New York: Wiley (1969).

J. Ledyard, "The pure theory of large two-candidate elections," Public Choice 44 (1984), 7-41.

I. Milchtaich, "Random-player games," Northwestern University discussion paper (1997).

P. Milgrom and R. Weber, "Distributional strategies for games with incomplete information," Mathematics of Operations Research 10 (1985), 619-632.

R. Myerson, "Population uncertainty and Poisson games," Northwestern University discussion paper (1994a).

R. Myerson, "Extended Poisson games and the Condorcet jury theorem," Northwestern University discussion paper (1994b), to appear in Games and Economic Behavior.

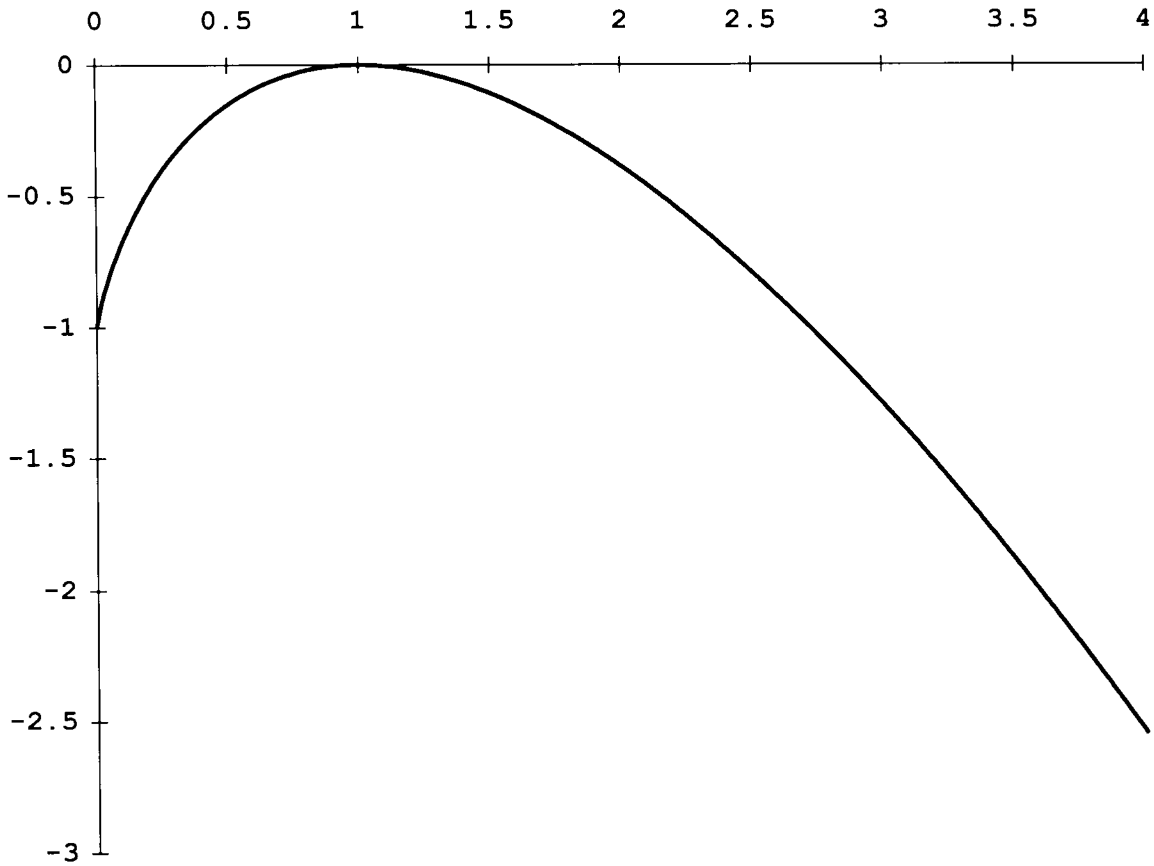

Figure 1. Graph of the Psi Function.