A Central Limit Theorem for Belief Functions*

Larry G. Epstein   Kyoungwon Seo

November 17, 2011

1. CLT for Belief Functions

The purpose of this Note is to prove a form of CLT (Theorem 1.4) that is used in Epstein and Seo (2011). More general central limit results and other applications will follow in later drafts.

Let $S = \{B, N\}$ and $\mathcal{K}(S) = \{\{B\}, \{N\}, \{B, N\}\}$ the set of nonempty subsets of S. Denote by $s^\infty = (s_1, s_2, ...)$ the generic element of S^∞ and by $\Psi_n(s^\infty)$ the empirical frequency of the outcome B in the first n experiments in sample s^∞. Let θ be a belief function on S, that is, there exists $m \in \Delta(\mathcal{K}(S))$ such that, for every $A \subset S$,

$$\theta(A) = m(\{K \in \mathcal{K}(S) : K \subset A\}).$$

Its conjugate θ^* is given by

$$\theta^*(A) = 1 - \theta(S \backslash A),$$

and the product θ^∞ is the belief function on S^∞ satisfying, for every $A \subset S^\infty$,

$$\theta^\infty(A) = m^\infty\left(\{\widetilde{K} = K_1 \times K_2 \times ... \in \mathcal{K}(S^\infty) : \widetilde{K} \subset A\}\right). \tag{1.1}$$

Here m^∞ is the ordinary i.i.d. product of the measure m.

---

*Boston University, lepstein@bu.edu and Department of Managerial Economics and Decision Sciences, Northwestern University, k-seo@kellogg.northwestern.edu. We gratefully acknowledge the financial support of the National Science Foundation (awards SES-0917740 and SES-0918248) and discussions with Zengjing Chen. This paper is preliminary; its main purpose is to provide a CLT that is used in our paper "Bayesian inference and non-Bayesian prediction and choice: foundations and an application to entry games with multiple equilibria."

The LLN asserts certainty that asymptotic empirical frequencies will lie in the interval $[\theta(B), 1-\theta(N)]$, that is,

$$\theta^{\infty}\{s^{\infty} : [\liminf \Psi_n(s^{\infty}), \limsup \Psi_n(s^{\infty})] \subset [\theta(B), 1-\theta(N)]\} = 1.$$

The CLT describes (up to approximation) beliefs about finite sample frequencies.

A simple CLT is provided first because it provides perspective on later results. Let $\boldsymbol{N}(\cdot)$ be the cdf of a standard normal distribution.

Theorem 1.1.

$$\begin{aligned}
\lim_{n\to\infty} \theta^{\infty}\left(\left\{s^{\infty} : \sqrt{n}\frac{\Psi_n(s^{\infty}) - \theta^*(B)}{\sqrt{\theta(\{N\})(1-\theta(\{N\}))}} \leq \alpha\right\}\right) &= \boldsymbol{N}(\alpha) \\
\lim_{n\to\infty} \theta^{\infty}\left(\left\{s^{\infty} : \sqrt{n}\frac{\Psi_n(s^{\infty}) - \theta(\{B\})}{\sqrt{\theta(\{B\})(1-\theta(\{B\}))}} > \alpha\right\}\right) &= 1 - \boldsymbol{N}(\alpha).
\end{aligned}$$

The proof follows readily from the next lemma showing that for the events indicated, the minimizing measures are i.i.d. As a result classical limit theorems applied to these measures deliver corresponding limit theorems for the i.i.d. product θ^{∞}.

Lemma 1.2. *Let P^* and P_* be the measures on S with $P_*(B) = \theta(B)$ and $P^*(B) = \theta^*(B)$ respectively. Then their i.i.d. products, denoted P_*^{∞} and $P^{*\infty}$, both lie in core (θ^{∞}); and, for any $0 \leq t \leq 1$,*

$$\theta^{\infty}(\{s^{\infty} : t \leq \Psi_n(s^{\infty})\}) = P_*^{\infty}(\{s^{\infty} : t \leq \Psi_n(s^{\infty})\}) \tag{1.2}$$

and

$$\theta^{\infty}(\{s^{\infty} : \Psi_n(s^{\infty}) \leq t\}) = P^{*\infty}(\{s^{\infty} : \Psi_n(s^{\infty}) \leq t\}). \tag{1.3}$$

See the appendix for a proof.

A CLT for two-sided intervals is less trivial because minimizing measures are not easily identified. Thus the following theorem uses a different proof strategy and applies a version of the multidimensional Berry-Esseen Theorem (Dasgupta (2008, pp. 145-6)): If the d-dimensional random variables $X_1, X_2, ...$ are i.i.d., $E(X_1) = 0$, $Var(X_1)$ is the identity matrix, and if $E(\|X_1\|)$ is finite, then there exists a constant K such that, for all n,

$$\sup_{C\in\mathcal{C}} \left| \Pr\left(\frac{X_1 + ... + X_n}{\sqrt{n}} \in C\right) - \Pr(Z \in C)\right| \leq \frac{K}{\sqrt{n}}.$$

Here $\mathcal{C}$ is the collection of all convex subsets of $\mathbb{R}^d$, and Z is standard normal and $\mathbb{R}^d$-valued.

Let $\boldsymbol{N}_2(\cdot,\cdot;\rho)$ be the cdf for the bivariate normal with zero means, unit variances and correlation coefficient ρ, that is,

$$\boldsymbol{N}_2(\alpha_1,\alpha_2;\rho) = \Pr(Z_1 \leq \alpha_1,\ Z_2 \leq \alpha_2)$$

where (Z_1, Z_2) is bivariate normal with the indicated moments.

Theorem 1.3. *There is a constant K that does not depend on α_1, α_2 or n, such that*

$$\left| \begin{array}{c} \theta^{\infty}\left(\left\{ \begin{array}{c} \theta(B) + \frac{\alpha_1}{\sqrt{n}}\sqrt{(1-\theta(B))\,\theta(B)} \leq \Psi_n(s^{\infty}) \\ \leq \theta^*(B) + \frac{\alpha_2}{\sqrt{n}}\sqrt{(1-\theta^*(B))\,\theta^*(B)} \end{array} \right\}\right) \\ -\boldsymbol{N}_2\left(-\alpha_1,\alpha_2;\frac{-\theta(B)\theta(N)}{\sqrt{\theta(B)(1-\theta(B))\theta^*(B)(1-\theta^*(B))}}\right) \end{array} \right| \leq \frac{K}{\sqrt{n}}.$$

Moreover, the same holds if α_1 and α_2 depend on n.

This theorem is a special case of the next one, but it also serves as a lemma in the proof of the more general result.

Remark 1. *When θ is additive, the indicated correlation coefficient ρ equals* 1 *and the inequality becomes*

$$\mid \theta^{\infty}\left(\alpha_1 < \sqrt{n}\tfrac{\Psi_n(s^{\infty})(B)-\theta(B)}{\sqrt{(1-\theta(B))\theta(B)}} \leq \alpha_2\right) - \Pr(\alpha_1 < Z \leq \alpha_2) \mid \leq \frac{K}{\sqrt{n}},$$

where Z is standard normal.

Proof. Define random variables[1]

$$X_i = I(K_i \in \mathcal{K}(S) : K_i \subset \{B\}) \text{ and } Y_i = 1 - I(K_i \in \mathcal{K}(S) : K_i \subset \{N\}). \quad (1.4)$$

[1]Note that $I(K_i \subset \{B\})) + I(K_i \subset \{N\})) = 1 - I(K_i \subset \{B,N\})) \neq 1$. Thus the first two indicators are not perfectly negatively correlated.

Observe first that $X_i \leq Y_i$. Compute, using m, that $E(X_i) = \theta(B)$, $E(Y_i) = 1 - \theta(N)$, $Var(X_i) = \theta(B)(1 - \theta(B))$, $Var(Y_i) = (1 - \theta(N))\theta(N)$ and

$$\begin{aligned} cov(X_i, Y_i) &= E(X_i Y_i) - E(X_i) E(Y_i) \\ &= E(X_i) - E(X_i) E(Y_i) \\ &= \theta(B) - \theta(B)(1 - \theta(N)) \\ &= \theta(B)\theta(N). \end{aligned}$$

Here, the second equality follows because $X_i = 1$ implies $Y_i = 1$. Then

$$\begin{aligned} E\left[\begin{pmatrix} \frac{X_i - \theta(B)}{\sqrt{(1-\theta(B))\theta(B)}} \\ \frac{Y_i - (1-\theta(N))}{\sqrt{(1-\theta(N))\theta(N)}} \end{pmatrix}\right] &= 0 \text{ and} \\ Var\left[\begin{pmatrix} \frac{X_i - \theta(B)}{\sqrt{(1-\theta(B))\theta(B)}} \\ \frac{Y_i - (1-\theta(N))}{\sqrt{(1-\theta(N))\theta(N)}} \end{pmatrix}\right] &= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \end{aligned}$$

where

$$\rho = corr(X_i, Y_i) = \frac{\theta(B)\theta(N)}{\sqrt{\theta(B)(1-\theta(B))(1-\theta(N))\theta(N)}}.$$

Note that

$$K_1 \times K_2 \times ... \subset \left\{ s^\infty : \beta_1 < \sum_{i=1}^n I(s_i = B) \leq \beta_2 \right\} \iff$$

$$\beta_1 < \min_{s^\infty \in K_1 \times K_2 \times ...} \sum_{i=1}^n I(s_i = B) \leq \max_{s^\infty \in K_1 \times K_2 \times ...} \sum_{i=1}^n I(s_i = B) \leq \beta_2 \iff$$

$$\beta_1 < \sum_{i=1}^n \min_{s^\infty \in K_1 \times K_2 \times ...} I(s_i = B) \leq \sum_{i=1}^n \max_{s^\infty \in K_1 \times K_2 \times ...} I(s_i = B) \leq \beta_2 \iff$$

$$\beta_1 < \sum_{i=1}^n I(K_i \subset \{B\}) \leq \sum_{i=1}^n [1 - I(K_i \subset \{N\})] \leq \beta_2 \iff$$

$$\beta_1 < \sum_{i=1}^n X_i \leq \sum_{i=1}^n Y_i \leq \beta_2$$

Conclude from (1.1) that

$$
\begin{aligned}
& \theta^{\infty}\left(\left\{s^{\infty}: \beta_1 < \sum_{i=1}^{n} I\left(s_i = B\right) \leq \beta_2\right\}\right) \\
= \;& m^{\infty}\left(\left\{K_1 \times K_2 \times \ldots \in \left(\mathcal{K}\left(S\right)\right)^{\infty}: \beta_1 < \sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} Y_i \leq \beta_2\right\}\right).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
& \theta^{\infty}\left(\alpha_1 < \tfrac{n\Psi_n(s^{\infty})-n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}}, \tfrac{n\Psi_n(s^{\infty})-n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \leq \alpha_2\right) \\
= \;& \theta^{\infty}\left(\begin{array}{c} \alpha_1\sqrt{n\left(1-\theta\left(B\right)\right)\theta\left(B\right)} + n\theta\left(B\right) < n\Psi_n\left(s^{\infty}\right) \\ \leq \sqrt{n\left(1-\theta\left(N\right)\right)\theta\left(N\right)}\alpha_2 + n\left(1-\theta\left(N\right)\right) \end{array}\right) \\
= \;& m^{\infty}\left(\begin{array}{c} \alpha_1\sqrt{n\left(1-\theta\left(B\right)\right)\theta\left(B\right)} + n\theta\left(B\right) < \sum_{i=1}^{n} X_i \\ \leq \sum_{i=1}^{n} Y_i \leq \sqrt{n\left(1-\theta\left(N\right)\right)\theta\left(N\right)}\alpha_2 + n\left(1-\theta\left(N\right)\right) \end{array}\right) \\
= \;& m^{\infty}\left(\alpha_1 < \tfrac{\sum_{i=1}^{n} X_i - n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}}, \tfrac{\sum_{i=1}^{n} Y_i - n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \leq \alpha_2\right).
\end{aligned}
$$

This permits translation of the assertion to be proven into one about i.i.d. probability measures and thus classical results can be applied.

Use the Cholesky decomposition of the variance-covariance matrix to obtain V^* such that $\left(V^*\right)^{-1}\begin{pmatrix} \frac{X_i-\theta(B)}{\sqrt{(1-\theta(B))\theta(B)}} \\ \frac{Y_i-(1-\theta(N))}{\sqrt{(1-\theta(N))\theta(N)}} \end{pmatrix}$ is standard normal (with correlation 0), and

$$
\left[\alpha_1 < \frac{\sum_{i=1}^{n} X_i - n\theta\left(B\right)}{\sqrt{n\left(1-\theta\left(B\right)\right)\theta\left(B\right)}} \text{ and } \frac{\sum_{i=1}^{n} Y_i - n\left(1-\theta\left(N\right)\right)}{\sqrt{n\left(1-\theta\left(N\right)\right)\theta\left(N\right)}} \leq \alpha_2\right] \iff
$$

$$
\left(V^*\right)^{-1}\begin{pmatrix} \frac{\sum_{i=1}^{n} X_i - n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}} \\ \frac{\sum_{i=1}^{n} Y_i - n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \end{pmatrix} \in C,
$$

for some convex $C \subset \mathbb{R}^2$. Therefore, by the multidimensional Berry-Esseen Theorem,

$$
\left| m^{\infty}\left(\left(V^*\right)^{-1}\begin{pmatrix} \frac{\sum_{i=1}^{n} X_i - n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}} \\ \frac{\sum_{i=1}^{n} Y_i - n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \end{pmatrix} \in C\right) - \Pr\left(\begin{pmatrix} Z \\ Z' \end{pmatrix} \in C\right)\right| \leq \frac{K}{\sqrt{n}}
$$

for some constant K that does not depend on C or n, where $\begin{pmatrix} Z \\ Z' \end{pmatrix}$ is standard normal.

Define

$$\begin{pmatrix} \tilde{Z} \\ \tilde{Z}' \end{pmatrix} = V^* \begin{pmatrix} Z \\ Z' \end{pmatrix}.$$

Then

$$\begin{aligned} \Pr\left(\begin{pmatrix} Z \\ Z' \end{pmatrix} \in C\right) &= \Pr\left(\alpha_1 < \tilde{Z}, \tilde{Z}' \le \alpha_2\right) \\ &= \Pr\left(-\tilde{Z} < -\alpha_1, \tilde{Z}' \le \alpha_2\right) \\ &= \boldsymbol{N}_2\left(-\alpha_1, \alpha_2; -\rho\right). \end{aligned}$$

Therefore,

$$\left| \begin{array}{c} \theta^\infty \left(\alpha_1 < \frac{n\Psi_n(s^\infty) - n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}}, \frac{n\Psi_n(s^\infty) - n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \le \alpha_2\right) \\ -\boldsymbol{N}_2\left(-\alpha_1, \alpha_2; \frac{-\theta(B)\theta(N)}{\sqrt{\theta(B)(1-\theta(B))(1-\theta(N))\theta(N)}}\right) \end{array} \right| \le \frac{K}{\sqrt{n}}.$$

Finally, the same proof works when α_1 and α_2 are replaced by $\alpha_{1,n}$ and $\alpha_{2,n}$.$\blacksquare$

The next theorem generalizes the preceding and is the main objective of this Note.

Theorem 1.4. *Suppose that $G : \mathbb{R} \to \mathbb{R}$ is bounded, quasi-concave and upper-semicontinuous. Then*

$$\int G\left(\Psi_n\left(s^\infty\right)\right) d\theta^\infty\left(s^\infty\right) = E\left[\min\left\{G\left(X'_{1n}\right), G\left(X'_{2n}\right)\right\}\right] + O\left(\frac{1}{\sqrt{n}}\right),$$

where (X'_{1n}, X'_{2n}) is normally distributed with mean $(\theta(B), \theta^*(B))$ and variance

$$\frac{1}{n}\begin{pmatrix} \theta(B)(1-\theta(B)) & \theta(B)\theta(N) \\ \theta(B)\theta(N) & (1-\theta(N))\theta(N) \end{pmatrix}.$$

That is,

$$\limsup_{n\to\infty} \sqrt{n} \left| \int G\left(\Psi_n\left(s^\infty\right)(B)\right) d\theta^\infty\left(s^\infty\right) - E\left[\min\left\{G\left(X'_{1n}\right), G\left(X'_{2n}\right)\right\}\right] \right| \le K$$

for some constant K. Moreover, the same holds when G depends on n and $\sup_{n,a} |G_n(a)| < \infty$.

The preceding theorem is the special case where $G(\cdot) = 1_{[a_n,b_n]}(\cdot)$ and

$$a_n = \theta(B) + n^{-1/2}\alpha_1\sqrt{(1-\theta(B))\,\theta(B)},\; b_n = \theta^*(B) + n^{-1/2}\alpha_2\sqrt{(1-\theta^*(B))\,\theta^*(B)}.$$

Proof. Without loss of generality, suppose that $\inf_{n,a} G(a) = 0$. Define $M = \sup_{n,a} G(a)$. Since G is quasiconcave, there are two (inverse) functions G_L^{-1} and G_R^{-1} such that

$$\{a : G(a) \geq t\} = \left[G_L^{-1}(t), G_R^{-1}(t)\right] \text{ for all } t \in [0, M].$$

(The inverses may not be defined at $t = 0$, but this is of no consequence below.)

By the definition of Choquet integration,

$$\begin{aligned}\int G(\Psi_n(s^\infty))\,d\theta^\infty(s^\infty) &= \int_0^M \theta^\infty(G(\Psi_n(s^\infty)) \geq t)\,dt \\ &= \int_0^M \theta^\infty\left(G_L^{-1}(t) \leq \Psi_n(s^\infty) \leq G_R^{-1}(t)\right)dt.\end{aligned}$$

Note that

$$\begin{aligned}&\theta^\infty\left(G_L^{-1}(t) \leq \Psi_n(s^\infty) \leq G_R^{-1}(t)\right) \\ = \;&\theta^\infty\begin{pmatrix} \frac{nG_L^{-1}(t)-n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}} \leq \frac{n\Psi_n(s^\infty)-n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}}, \\ \frac{n\Psi_n(s^\infty)-n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \leq \frac{nG_R^{-1}(t)-n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}} \end{pmatrix}.\end{aligned}$$

Thus, by Theorem 1.3,

$$\begin{aligned}&\theta^\infty\left(G_L^{-1}(t) \leq \Psi_n(s^\infty) \leq G_R^{-1}(t)\right) \\ = \;&\boldsymbol{N}_2\left(-\frac{nG_L^{-1}(t)-n\theta(B)}{\sqrt{n(1-\theta(B))\,\theta(B)}}, \frac{nG_R^{-1}(t)-n(1-\theta(N))}{\sqrt{n(1-\theta(N))\,\theta(N)}}; -\rho\right) + O\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

with $\rho = \frac{\theta(B)\theta(N)}{\sqrt{\theta(B)(1-\theta(B))(1-\theta(N))\theta(N)}}$. Because the term $O\left(\frac{1}{\sqrt{n}}\right)$ does not depend on t,

$$\begin{aligned}&\int_0^M \theta^\infty\left(G_L^{-1}(t) \leq \Psi_n(s^\infty) \leq G_R^{-1}(t)\right)dt \\ = \;&\int_0^M \boldsymbol{N}_2\left(-\frac{nG_L^{-1}(t)-n\theta(B)}{\sqrt{n(1-\theta(B))\,\theta(B)}}, \frac{nG_R^{-1}(t)-n(1-\theta(N))}{\sqrt{n(1-\theta(N))\,\theta(N)}}; -\rho\right)dt + O\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Let Z_1 and Z_2 be jointly normally distributed with mean $(0,0)$ and variance $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, and let X'_{1n} and X'_{2n} be normally distributed as in the theorem statement. Then

$$
\begin{aligned}
& \int_0^M \boldsymbol{N}_2 \left(-\frac{nG_L^{-1}(t) - n\theta(B)}{\sqrt{n(1-\theta(B))\theta(B)}}, \frac{nG_R^{-1}(t) - n(1-\theta(N))}{\sqrt{n(1-\theta(N))\theta(N)}}; -\rho \right) dt \\
= & \int_0^M \Pr \begin{pmatrix} G_L^{-1}(t) \le \theta(B) + \frac{\sqrt{n(1-\theta(B))\theta(B)}}{n} Z_1, \\ (1-\theta(N)) + \frac{\sqrt{n(1-\theta(N))\theta(N)}}{n} Z_2 \le G_R^{-1}(t) \end{pmatrix} dt \\
= & \int_0^M \Pr \left(G_L^{-1}(t) \le X'_{1n},\, X'_{2n} \le G_R^{-1}(t) \right) dt \\
\overset{*}{=} & \int_0^M \Pr \left(G_L^{-1}(t) \le X'_{1n} \le G_R^{-1}(t),\, G_L^{-1}(t) \le X'_{2n} \le G_R^{-1}(t) \right) dt + O\left(\frac{1}{\sqrt{n}}\right) \\
= & \int_0^M \Pr \left(G(X'_{1n}) \ge t,\, G(X'_{2n}) \ge t \right) dt + O\left(\frac{1}{\sqrt{n}}\right) \\
= & \int_0^M \Pr \left(\min\{G(X'_{1n}), G(X'_{2n})\} \ge t \right) dt + O\left(\frac{1}{\sqrt{n}}\right) \\
= & E\left[\min\{G(X'_{1n}), G(X'_{2n})\}\right] + O\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

To complete the proof, we need only prove the equality marked with an asterisk. Define random variables $(X_i, Y_i)_{i=1}^{\infty}$ as in (1.4). Then they are i.i.d. under m^{∞} and, because $X_i \le Y_i$,

$$
\begin{aligned}
& \Pr\left(a \le \frac{\sum_{i=1}^n X_i}{n}, \frac{\sum_{i=1}^n Y_i}{n} \le b \right) \\
= & \Pr\left(a \le \frac{\sum_{i=1}^n X_i}{n} \le \frac{\sum_{i=1}^n Y_i}{n} \le b \right) \\
= & \Pr\left(a \le \frac{\sum_{i=1}^n X_i}{n} \le b,\, a \le \frac{\sum_{i=1}^n Y_i}{n} \le b \right).
\end{aligned}
$$

By the multidimensional Berry-Esseen Theorem,

$$
\begin{aligned}
& \left| \quad \Pr\left(a \leq \frac{\sum_{i=1}^{n} X_i}{n},\, \frac{\sum_{i=1}^{n} Y_i}{n} \leq b\right) - \Pr(a \leq X'_{1n},\, X'_{2n} \leq b) \right| \\
= & \left| \Pr\left((a - E[X_1])\sqrt{n} \leq \frac{\sum_{i=1}^{n}(X_i - E[X_1])}{\sqrt{n}},\, \frac{\sum_{i=1}^{n}(Y_i - E[Y_1])}{\sqrt{n}} \leq (b - E[Y_1])\sqrt{n}\right) \right. \\
& \left. - \Pr\left((a - E[X_1])\sqrt{n} \leq (X'_{1n} - E[X_1])\sqrt{n},\; (X'_{2n} - E[Y_1])\sqrt{n} \leq (b - E[Y_1])\sqrt{n}\right) \right| \\
= & \sup_{C \in \mathbb{R}^2 \text{ convex}} \left| \Pr\left(\left(\frac{\sum_{i=1}^{n}(X_i - E[X_1])}{\sqrt{n}},\, \frac{\sum_{i=1}^{n}(Y_i - E[Y_1])}{\sqrt{n}}\right) \in C\right) - \Pr((Z_1, Z_2) \in C) \right| \leq \frac{K}{\sqrt{n}},
\end{aligned}
$$

and similarly

$$
\left| \Pr\left(a \leq \frac{\sum_{i=1}^{n} X_i}{n} \leq b,\, a \leq \frac{\sum_{i=1}^{n} Y_i}{n} \leq b\right) - \Pr\left(a \leq X'_{1n} \leq b,\, a \leq X'_{2n} \leq b\right) \right| \leq \frac{K}{\sqrt{n}}.
$$

It follows that, for all $a \leq b$ and n,

$$
\left| \Pr(a \leq X'_{1n},\, X'_{2n} \leq b) - \Pr\left(a \leq X'_{1n} \leq b,\, a \leq X'_{2n} \leq b\right) \right| \leq \frac{2K}{\sqrt{n}}.
$$

This proves the marked equation.

Finally, the above proof works also if G is replaced by G_n such that $\sup_{n,a} |G_n(a)| \leq \infty$. ■

A. Appendix: Proof of Lemma 1.2

Prove (1.2).
Step 1: $K_1 \times K_2 \times ... \subset \{s^\infty : \Sigma_{i=1}^{n} I(s_i = B) \geq t\}$ iff $\Sigma_{i=1}^{n} I(K_i \subset \{B\}) \geq t$.
Here is a proof:

$$
\begin{aligned}
K_1 \times K_2 \times ... \quad & \subset \quad \left\{s^\infty : \sum_{i=1}^{n} I(s_i = B) \geq t\right\} \\
& \iff \quad \min_{s^\infty \in K_1 \times K_2 \times ...} \sum_{i=1}^{n} I(s_i = B) \geq t \\
& \iff \quad \sum_{i=1}^{n} \min_{s_i \in K_i} I(s_i = B) \geq t \\
& \iff \quad \sum_{i=1}^{n} I(K_i \subset \{B\}) \geq t.
\end{aligned}
$$

Step 2: Let $m^{\infty} \in \Delta(\mathcal{K}(S^{\infty}))$ be the measure for θ^{∞}. Then

$$\begin{aligned} & \theta^{\infty}\left(\left\{s^{\infty} : \sum_{i=1}^{n} I(s_i = B) \geq t\right\}\right) \\ = \; & m^{\infty}\left(\left\{K_1 \times K_2 \times ... \in \mathcal{K}(S^{\infty}) : \sum_{i=1}^{n} I(K_i \subset \{B\}) \geq t\right\}\right). \end{aligned}$$

Argue as follows:

$$\begin{aligned} & \theta^{\infty}\left(\left\{s^{\infty} : \sum_{i=1}^{n} I(s_i = B) \geq t\right\}\right) \\ = \; & m^{\infty}\left(\left\{K \in \mathcal{K}(S^{\infty}) : K \subset \left\{s^{\infty} : \sum_{i=1}^{n} I(s_i = B) \geq t\right\}\right\}\right) \\ = \; & m^{\infty}\left(\left\{K_1 \times K_2 \times ... \in \mathcal{K}(S^{\infty}) : K_1 \times K_2 \times ... \subset \left\{s^{\infty} : \sum_{i=1}^{n} I(s_i = B) \geq t\right\}\right\}\right). \end{aligned}$$

Next apply Step 1.
Step 3: Complete the proof. By Step 2,

$$\begin{aligned} & \theta^{\infty}(\{s^{\infty} : \Psi_n(s^{\infty}) \geq t\}) \\ = \; & \theta^{\infty}\left(\left\{s^{\infty} : \tfrac{1}{n}\sum_{i=1}^{n} I(s_i = B) \geq t\right\}\right) \\ = \; & m^{\infty}\left(\left\{K_1 \times K_2 \times ... \in \mathcal{K}(S^{\infty}) : \tfrac{1}{n}\sum_{i=1}^{n} I(K_i = \{B\}) \geq t\right\}\right) \\ = \; & \Pr\left(\tfrac{1}{n}\Sigma_{i=1}^{n} Y_i \geq t\right), \end{aligned}$$

where $Y_i = 0$ or 1, $\Pr(Y_i = 1) = \theta(B)$ and the Y_i's are i.i.d. Therefore, the preceding equals $P_*^{\infty}(\{s^{\infty} : \Psi_n(s^{\infty}) \geq nt\})$.

To prove (1.3), reverse the roles of B and N in the preceding argument. ■

References

[1] A. Dasgupta, *Asymptotic Theory of Statistics and Probability*, Springer, 2008.

[2] L.G. Epstein and K. Seo, Bayesian inference and non-Bayesian prediction and choice: foundations and an application to entry games with multiple equilibria, 2011.