## 1.5. LIMIT DISTRIBUTIONS AND THE CONTINUITY THEOREM.

Let $\{X_n, n \geq 0\}$ be non-negative, integer valued random variables with $(n \geq 0, k \geq 0)$

$$P[X_n = k] = p_k^{(n)}, \quad P_n(s) = Es^{X_n}. \tag{1.5.1}$$

Then $X_n$ converges in distribution to $X_0$, written $X_n \Rightarrow X_0$, if

$$\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)} \tag{1.5.2}$$

for $k = 0, 1, 2, \ldots$. As the next result shows, this is equivalent to

$$P_n(s) \to P_0(s) \tag{1.5.3}$$

for $0 \leq s \leq 1$ as $n \to \infty$.

**Theorem 1.5.1. The Continuity Theorem.** *Suppose for each $n = 1, 2, \ldots$ that $\{p_k^{(n)}, k \geq 0\}$ is a probability mass function on $\{0, 1, 2, \ldots\}$ so that*

$$p_k^{(n)} \geq 0, \quad \sum_{k=0}^{\infty} p_k^{(n)} = 1.$$

*Then there exists a sequence $\{p_k^{(0)}, k \geq 0\}$ such that*

$$\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)}, \quad k \geq 0, \tag{1.5.4}$$

*iff there exists a function $P_0(s), 0 < s < 1$ such that*

$$\lim_{n\to\infty} P_n(s) = \lim_{n\to\infty} \sum_{k=0}^{\infty} p_k^{(n)} s^k = P_0(s). \tag{1.5.5}$$

*for $0 < s < 1$. In this case $P_0(s) = \sum_{k=0}^{\infty} p_k^{(0)} s^k$ and $\sum_{k=0}^{\infty} p_k^{(0)} = 1$ iff $\lim_{s\uparrow 1} P_0(s) =: P_0(1) = 1$.*

**Remarks.** As we will see, this provides an alternative to brute force when proving convergence in distribution. Frequently it is easier to prove that the generating functions converge rather than trying to show the convergence of a sequence of mass functions.

From (1.5.4) we have

$$0 \leq p_k^{(0)} \leq 1 \tag{1.5.6}$$

since the same is true for $p_k^{(n)}$ and $\lim_{n\to\infty} p_k^{(n)} = p_k^{(0)}$. But it does not follow that $\sum_{k=0}^{\infty} p_k^{(0)} = 1$ since mass can escape to infinity. As a graphic example suppose

$$p_k^{(n)} = \delta_{k,n} = \begin{cases} 1, & \text{if } k = n \\ 0, & \text{if } k \neq n. \end{cases}$$

For any fixed $k$,

$$\lim_{n\to\infty} p_k^{(n)} = 0, \tag{1.5.7}$$

from which

$$(p_0^{(0)}, p_1^{(0)}, \dots) = (0, 0, \dots).$$

This phenomenon arises because we consider the state space $\{0, 1, 2, \dots\}$. If we enlarge the state space to $\{0, 1, 2, \dots, \infty\}$ then $X_n \Rightarrow \infty$ and the limit distribution concentrates all mass at $\infty$.

*Proof.* Suppose (1.5.2). Fix $s \in (0,1)$ and for any $\epsilon > 0$ we may pick $m$ so large that

$$\sum_{i=m+1}^{\infty} s^i < \epsilon.$$

We have

$$\begin{aligned} |P_n(s) - P_0(s)| &\le \sum_{k=1}^{\infty} |p_k^{(n)} - p_k^{(0)}| s^k \\ &\le \sum_{k=1}^{m} |p_k^{(n)} - p_k^{(0)}| + \sum_{k=m+1}^{\infty} s^k \\ &\le \sum_{k=1}^{m} |p_k^{(n)} - p_k^{(0)}| + \epsilon. \end{aligned}$$

Letting $n \to \infty$, we get

$$\limsup_{n\to\infty} |P_n(s) - P_0(s)| \le \epsilon$$

and because $\epsilon$ is arbitrary we obtain (1.5.5).

The proof of the converse is somewhat more involved and is deferred to the appendix at the end of this section, which can be read by the interested student or skipped by a beginner.



**Example. Harry and the Mushroom Staples.*** Each morning, Harry buys enough salad ingredients to prepare 200 salads for the lunch crowd at his restaurant. Included in the salad are mushrooms which come in small boxes held shut by relatively large staples. For each salad, there is probability .005 that the person preparing the salad will sloppily drop a staple into it. During a three week period, Harry's precocious twelfth grade niece, who has just completed a statistics unit in high school, keeps track of the number of staples dropped in salads. (Harry's customers are not reticent about complaining about such things so detection of the sin and collection of the data pose no problem.) After drawing a histogram, the niece decides that the number of salads per day containing a staple is Poisson distributed with parameter $(200)(.005) = 1$. ■

Harry's niece has empirically rediscovered the Poisson approximation to the binomial distribution: If $X_n \sim b(k; n, p(n))$ and

$$\lim_{n\to\infty} np(n) = \lim_{n\to\infty} EX_n = \lambda \in (0, \infty), \tag{1.5.8}$$

then

$$X_n \Rightarrow X_0$$

as $n \to \infty$ where $X_0 \sim p(k; \lambda)$.

The verification is easy using generating functions. We have

$$\begin{aligned} \lim_{n\to\infty} P_n(s) = \lim_{n\to\infty} Es^{X_n} &= \lim_{n\to\infty} (1 - p(n) + p(n)s)^n \\ &= \lim_{n\to\infty} \left(1 + \frac{(s-1)np(n)}{n}\right)^n = e^{\lambda(s-1)} \end{aligned}$$

using (1.5.8).

**Appendix: Continuation of the Proof of Theorem 1.5.1.**

We now return to the proof of Theorem 1.5.1 and show why convergence of the generating functions implies convergence of the sequences.

Assume we know the following fact: Any sequence of mass functions $\{\{f_j^{(n)}, j \ge 0\}, n \ge 1\}$ has a convergent subsequence $\{\{f_j^{(n')}, j \ge 0\}\}$ meaning that for all $j$

$$\lim_{n'\to\infty} f_j^{(n')}$$

exists. If $\{p_k^{(n)}\}$ has two different subsequential limits along $\{n'\}$ and $\{n''\}$, by the first half of Theorem 1.5.1 and hypothesis (1.5.3), we would have

$$\lim_{n'\to\infty} \sum_{k=0}^{\infty} p_k^{(n')} s^k = \lim_{n'\to\infty} P_{n'}(s) = P_0(s)$$

*A semi-true story.

and also

$$\lim_{n''\to\infty} \sum_{k=0}^{\infty} p_k^{(n'')} s^k = \lim_{n''\to\infty} P_{n''}(s) = P_0(s).$$

Thus any two subsequential limits of $\{p_k^{(n)}\}$ have the same generating function. Since generating functions uniquely determine the sequence, all subsequential limits are equal and thus $\lim_{n\to\infty} p_k^{(n)}$ exists for all $k$. The limit has a generating function $P_0(s)$.

It remains to verify the claim that a sequence of mass functions $\{\{f_j^{(n)}, j \geq 0\},\ n \geq 1\}$ has a subsequential limit. Since for each $n$ we have

$$\{f_j^{(n)}, j \geq 0\} \subset [0,1]^\infty,$$

and $[0,1]^\infty$ is a compact set (being a product of the compact sets $[0,1]$), we have an infinite sequence of elements in a compact set. Hence a subsequential limit must exist.

If the compactness argument is not satisfying, a subsequential limit can be manufactured by a diagonalization procedure. (See Billingsley, 1986, page 566.)

### 1.5.1. The Law of Rare Events.

A more sophisticated version of the Poisson approximation, sometimes called the Law of Rare Events, is discussed next.

**Proposition 1.5.2.** *Suppose we have a doubly indexed array of random variables such that for each $n = 1, 2, \ldots$, $\{X_{n,k}, k \geq 1\}$, is a sequence of independent (but not necessarily identically distributed) Bernoulli random variables satisfying*

$$P[X_{n,k} = 1] = p_k(n) = 1 - P[X_{n,k} = 0], \tag{1.5.1.1}$$

$$\bigvee_{1\leq k\leq n} p_k(n) =: \delta(n) \to 0, \quad n \to \infty, \tag{1.5.1.2}$$

$$\sum_{k=1}^{n} p_k(n) = E\sum_{k=1}^{n} X_{n,k} \to \lambda \in (0,\infty), \quad n \to \infty. \tag{1.5.1.3}$$

*If $PO(\lambda)$ is a Poisson distributed random variable with mean $\lambda$ then*

$$\sum_{k=1}^{n} X_{n,k} \Rightarrow PO(\lambda).$$