DEAR 310A STUDENTS; ON OUR NEXT TO LAST HOMEWORK (ON STEIN'S METHOD) I ASKED A VAGUELY WORDED QUESTION ABOUT THE BIRTHDAY PROBLEM WITH UNEQUAL BOX PROBABILITIES. I ASSUMED THAT IF $\binom{n}{2}\sum_i p_i^2 \to \lambda$ THEN $P\{\text{NO MATCH}\} \sim e^{-\lambda}$. SEVERAL OF YOU NOTICED THAT THIS DOES <u>NOT</u> FOLLOW FROM OUR BOUND (STEIN'S METHOD) AND SOME EXTRA HYPOTHESIS ON THE $\{p_i\}_i$ IS NEEDED. THEY ARE RIGHT!

THE SITUATION IS A BIT COMPLICATED BUT A REASONABLY DEFINITIVE ANSWER CAN BE FOUND IN: CAMARRI, M. AND PITMAN, J. (2000) LIMIT DISTRIBUTIONS AND RANDOM TREES DERIVED FROM THE BIRTHDAY PROBLEM WITH UNEQUAL PROBABILITIES. ELECTRONIC JOUR. PROBAB. VOL 5, #2, PG 1-18. HERE IS A BRIEF DESCRIPTION/DISCUSSION.

FOR EACH n, LET $p_{n1} \ge p_{n2} \ge \cdots$ BE A PROBABILITY; p_{ni} = CHANCE OF BOX i IN n^{TH} ROW OF A TRIANGULAR ARRAY. LET (FOR THE n^{th} ROW), BALLS BE DROPPED INTO BOXES. LET R_n BE THE FIRST REPEAT TIME. THE CLASSICAL BIRTHDAY PROBLEM HAS $p_{n,i} = 1/n$, $1 \le i \le n$, WE KNOW $P\{R_n > c\sqrt{n}\} = P\{\text{IF } c\sqrt{n} \text{ BALLS DROPPED INTO } n \text{ BOXES, NO MATCH}\} \sim e^{-\lambda}$ WITH $\lambda = \binom{c\sqrt{n}}{2}\frac{1}{n} \sim \frac{c^2}{2}$, SO $P\{R_n > c\sqrt{n}\} \sim e^{-c^2/2}$. HERE IS THEIR RESULT

<u>THEOREM</u> LET $s_n = \sqrt{\sum_i p_{ni}^2}$ AND $\theta_{ni} = p_{ni}/s_n$.

(1) IF $p_{n1} \to 0$ AS $n \to \infty$ AND $\theta_i = \lim_n \theta_{ni}$ EXISTS FOR EACH i, THEN, FOR ANY $c \ge 0$

$$\lim_{n\to\infty} P\{R_n > c/s_n\} = e^{-\frac{1}{2}(1-\sum \theta_i^2)c^2} \prod_i (1+\theta_i c)e^{-\theta_i c}.$$

(2) CONVERSELY, IF THERE ARE POSITIVE CONSTANTS $c_n \to 0$, d_n SUCH THAT $c_n(R_n - d_n)$ HAS A NON DEGENERATE LIMIT, THEN $p_{n1} \to 0$, THE LIMITS θ_i EXIST AND SO THE LIMIT IS AS ABOVE WITH $c_n/s_n \to \alpha$ FOR SOME $0 < \alpha < \infty$ AND $c_n d_n \to 0$.

<u>EXAMPLES</u> (a) IF $p_{ni} = 1/n$, $1 \le i \le n$, $s_n = 1/\sqrt{n}$, $\theta_i \equiv 0$ AND $P\{R_n > c\sqrt{n}\} \to e^{-c^2/2}$.

(b) IF $p_{ni} = \frac{1}{i H_n}$, $1 \le i \le n$, $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \sim \log n$, $s_n \sim \log n \sqrt{\pi^2/6}$, $\theta_i = 1/i\sqrt{\pi^2/6}$, $1 \le i < \infty$ AND NOTHING LIKE A POISSON APPROXIMATION HOLDS!

SORRY, <u>BUT</u> THAT'S WHY WE PROVE THEOREMS!

P.S. YOU MAY CHECK (PLEASE DO) THAT IF $\binom{n}{2}\sum p_i^2 \to \lambda$ <u>AND</u> $p_{n1}/\sqrt{\sum p_{ni}^2} \to 0$ THEN POISSON APPROX O.K.