The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

1863-6

Advanced School and Conference on Statistics and Applied Probability in Life Sciences

24 September - 12 October, 2007

Poisson Approximation

Louis Chen

National University of Singapore
Institute for Mathematical Sciences
Singapore

POISSON APPROXIMATION

A series of 14 one-hour lectures
given in January 2003 (Semester 2, 2002-03)
at the Dept of Maths., NUS, as
part of the module MA 6291
Special Topics in Mathematics I:
Probability approximations.

References

1. Richard Arratia, Larry Goldstein and Louis Gordon (1990), Poisson approximation and the Chen-Stein method, *Statistical Science* **5**, 403-434

2. A. D. Barbour, Lars Holst and Svante Janson (1992), *Poisson Approximation*, Oxford Studies in Probability **2**, Clarendon Press, Oxford

3. Torkel Erhardsson (2005). Stein's method for Poisson and compound Poisson approximation, *An Introduction to Stein's Method* (A.D. Barbour and L. H. Y. Chen, eds), Lecture Notes Series **No. 4**, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific, 61-113

1.

Poisson approximation

§1 Classical Poisson limit theorem

Siméon-Denis Poisson (1837):

$S_n \sim Bi(n,p)$

$P(S_n = k) \to e^{-\lambda}\frac{\lambda^k}{k!}$ for $k=0,1,\dots$

as $n\to\infty$ & $p\to 0$ s.t. $np\to\lambda$ $(0<\lambda<\infty)$

$P(S_n=k) = \binom{n}{k}p^k(1-p)^{n-k}$

So can prove using calculus.

Theorem 1.1 $X_{n1},\dots,X_{nn}$ indep

$P(X_{ni}=1) = 1-P(X_{ni}=0) = p_{ni}$

$\tilde{p}_n = \max_{1\le i\le n} p_{ni} \to 0$, $\lambda_n = \sum_{i=1}^n p_{ni} \to \lambda > 0$

as $n\to\infty$.

$W_n = \sum_{i=1}^n X_{ni}$, $Z\sim Po(\lambda)$

i.e. $P(Z=k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $k=0,1,\dots$

then $\forall$ $k=0,1,2,\dots$

$P(W_n=k)\to P(Z=k)$ as $n\to\infty$.

<u>Proof.</u> $E W_n f(W_n) = \sum_{i=1}^{n} E X_{ni} f(W_n)$

$$= \sum_{i=1}^{n} p_{ni} E f(W_n^{(i)}+1) \quad \text{where } W_n^{(i)} = W_n - X_{ni}$$

$$= \sum_{i=1}^{n} p_{ni} E[f(W_n^{(i)}+1) - f(W_n+1)]$$

$$+ \lambda_n E f(W_n+1)$$

$$= \lambda_n E f(W_n+1) + \sum_{i=1}^{n} p_{ni}^2 E[f(W_n^{(i)}+1) - f(W_n^{(i)}+2)]$$

Now $P(W_n = 0) = \prod_{i=1}^{n}(1-p_{ni}) = e^{\log \prod_{i=1}^{n}(1-p_{ni})}$

$$= e^{\sum_{i=1}^{n}\left(-p_{ni} - \frac{p_{ni}^2}{2} - \frac{p_{ni}^3}{3} - \cdots\right)}$$

$$= e^{-\lambda_n + O\left(\lambda_n\left(1 + \frac{\tilde{p}_n}{2} + \frac{\tilde{p}_n^2}{3} + \cdots\right)\right)}$$

$$\to e^{-\lambda}$$

For $k = 1, 2, \ldots$, let $f = I_{\{k\}}$

then $kP(W_n = k) = \lambda_n P(W_n = k-1)$

$$\underbrace{- \sum_{i=1}^{n} p_{ni}^2 E[\Delta f(W_n^{(i)}+1)]}_{\downarrow \; 0}$$

Since $P(W_n=0)$ has a positive limit,

$P(W_n=k)$ also have a positive limit for $k\geq 1$

by induction.

Let $p_k = \lim_{n\to\infty} P(W_n=k)$ for $k=0,1,\ldots$,

then for $k\geq 1$,

$$k p_k = \lambda p_{k-1}.$$

Solving this difference equation using $p_0 = e^{-\lambda}$

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{for } k=0,1,2,\ldots$$

Hence $P(W_n=k) \to P(Z=k)$ as $n\to\infty$.

<u>Corollary 1.2</u>

$$\sum_{k=0}^{\infty} |P(W_n=k) - P(Z=k)| \to 0 \text{ as } n\to\infty$$

<u>Proof.</u> $\sum_{k=0}^{\infty} |P(W_n=k) - P(Z=k)|$

$$= \sum_{k=0}^{N} |P(W_n=k) - P(Z=k)| + P(W_n>N) + P(Z>N)$$

$$\leq \sum_{k=0}^{N} |P(W_n=k) - P(Z=k)| + \frac{EW_n}{N} + \frac{EZ}{N}$$

$$= \sum_{k=0}^{N} |P(X_n = k) - P(Z = k)| + \frac{\lambda_1}{N} + \frac{\lambda}{N}$$

$\to 0$ by letting $n \to \infty$ & then $N \to \infty$.

§2 Total variation distance

P, Q prob. measures on $(\Omega, \mathcal{B})$

Define total variation distance $d_{TV}(P, Q)$ (or simply $d(P,Q)$) between P & Q by

$$d(P,Q) = \sup_{A \in \mathcal{B}} |P(A) - Q(A)|$$

Define the norm $\|P-Q\|$ of the signed measure $P-Q$ (or of the total mass of the total variation $|P-Q|$) by

$$\|P-Q\| = \sup_{|h|=1} \left| \int h\,dP - \int h\,dQ \right|$$

where h is measurable

Fact

(1) Both $d(\cdot,\cdot)$ and $\|\cdot - \cdot\|$ are metrics.

(2) $\|P-Q\| = \sup_{|h| \le 1} \left| \int h\,dP - \int h\,dQ \right|$

(3) $\|P-Q\| = P(\Delta) - Q(\Delta) - [P(\Delta^c) - Q(\Delta^c)]$

where Δ is a maximal set s.t. $P(\Delta) \geq Q(\Delta)$.

$$= [P(\Delta) - Q(\Delta)] - [1 - P(\Delta) + 1 - Q(\Delta)]$$

$$= 2[P(\Delta) - Q(\Delta)]$$

$$= 2d(P,Q)$$

(4) $d(P,Q) = \sup_{0 \leq h \leq 1} \left|\int h\,dP - \int h\,dQ\right|$

Let X & Y be random variables with distributions denoted by $\mathcal{L}(X)$, $\mathcal{L}(Y)$.

Then $\|\mathcal{L}(X) - \mathcal{L}(Y)\| = \sup_{|h| \leq 1} |Eh(X) - Eh(Y)|$

& $d(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \in \mathcal{B}(\mathbb{R})} |P(X \in A) - P(Y \in A)|$

If X & Y are integer-valued, then

$$\|\mathcal{L}(X) - \mathcal{L}(Y)\| = \sum_{k \in \mathbb{Z}} |P(X=k) - P(Y=k)|$$

$$d(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subset \mathbb{Z}} |P(X \in A) - P(Y \in A)|$$

Corollary 1.2 $\Rightarrow$ $d(\mathcal{L}(W_n), \mathcal{L}(Z)) \to 0$

as $n \to \infty$.

~~What is the rate of convergence or how close $\mathcal{L}(W_n)$ is to $\mathcal{L}(Z)$?~~

Now drop the subscript n for simplicity

$X_1, \ldots, X_n$ indep. $P(X_i = 1) = 1 - P(X_i = 0) = p_i$

$\lambda = \sum_{i=1}^{n} p_i$, $W = \sum_{i=1}^{n} X_i$, $Z \sim Po(\lambda)$, $\tilde{p} = \max_{1 \le i \le n} p_i$

(1) If $p_1 = \cdots = p_n = p$, then $\lambda = np$ & $W \sim B(n,p)$.

Prokhorov (Uspehi Mat. Nauk (1953)):

$$d(Bi(n,p), Po(\lambda)) \le p\left[\frac{1}{\sqrt{2\pi e}} + O\left(\min\left(1, \frac{1}{\sqrt{\lambda}} + p\right)\right)\right]$$

For the following the p_i are not necessarily equal.

(2) Hodges & LeCam (Ann. Math. Stat. (1960)):

$$\max_{z \ge 0} |P(W \le z) - P(Z \le z)| \le 3\tilde{p}^{\frac{1}{3}}.$$

(3) LeCam (Pacific J. Math (1960)):

(i) $d(\mathcal{L}(W), \mathcal{L}(Z)) \leq \sum_{i=1}^{n} p_i^2$

(ii) $d(\mathcal{L}(W), \mathcal{L}(Z)) \leq 4.5\,\tilde{p}$

(iii) $d(\mathcal{L}(W), \mathcal{L}(Z)) \leq 8\,\frac{1}{\lambda} \sum_{i=1}^{n} p_i^2 \leq 8\tilde{p}$

$(\tilde{p} \leq \frac{1}{4})$

(proof using convolution operators)

(4) Kerstan (ZW (1964)):

$d(\mathcal{L}(W), \mathcal{L}(Z)) \leq 1.05\,\frac{1}{\lambda} \sum_{i=1}^{n} p_i^2$

$(\tilde{p} \leq \frac{1}{4})$

(proof using characteristic function)

Proof of 3(i).

Let $\xi_1, \dots, \xi_n$ be indep & indep of $X_1, \dots, X_n$.

$\xi_i \sim Po(p_i)$. Then $\mathcal{L}(Z) = \mathcal{L}(\sum_{i=1}^{n} \xi_i)$

$d(\mathcal{L}(W), \mathcal{L}(Z)) = d(\mathcal{L}(W), \mathcal{L}(\sum_{i=1}^{n} \xi_i))$

No

Date

$\mathcal{J}_c$

$$\le \sum_{i=1}^{n} d\Big(\mathcal{L}(\xi_1+\dots+\xi_{i-1}+X_i+\dots+X_n),$$

$$\mathcal{L}(\xi_1+\dots+\xi_i+X_{i+1}+\dots+X_n)\Big)$$

$$\le \sum_{i=1}^{n} d(\mathcal{L}(X_i), \mathcal{L}(\xi_i))$$

$$= \sum_{i=1}^{n} \left[P(X_i=1) - P(\xi_i=1)\right]$$

$$= \sum_{i=1}^{n} (p_i - p_i e^{-p_i}) = \sum_{i=1}^{n} p_i(1-e^{-p_i})$$

$$\le \sum_{i=1}^{n} p_i^2 .$$

It is far from trivial to have the factor $\frac{1}{\lambda}$ in the bound.

With $\frac{1}{\lambda}$ as a factor, the bound $\to 0$ as $\tilde{p} \to 0$, irrespective of λ.

whereas $\sum_{i=1}^{n} p_i^2 = n \cdot \frac{1}{n} = 1 \not\to 0$

if $p_i = \frac{1}{\sqrt{n}}$.

The method of Le Cam & Kerstan work well only for indep. r.v.'s. A method due to Stein (1972) & Chen (1975) works much better for dependent r.v.'s.

§3 The Chen-Stein method

Stein (6th Berkeley Symposium (1972)): Normal approximation

Chen (Ann. Prob. (1975)): Poisson approximation.

Characterization of the Poisson distribution

Proposition 3.1 $Z \sim Po(\lambda)$ if and only if $\forall$ bdd $f: \mathbb{Z}^+ \to \mathbb{R}$,

$$E\{\lambda f(Z+1) - Z f(Z)\} = 0 \quad (3.1).$$

Proof. If $Z \sim Po(\lambda)$, then

$$E Z f(Z) = \sum_{k=0}^{\infty} k f(k) e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= \sum_{k=1}^{\infty} f(k) \cdot e^{-\lambda} \frac{\lambda^k}{(k-1)!} = \sum_{l=0}^{\infty} f(l+1) e^{-\lambda} \frac{\lambda^{l+1}}{l!}$$

$$= \lambda E f(Z+1) \Rightarrow (3.1)$$

If (3.1) holds, then by letting $f(k) = I_{\{k\}}$ for $k = 1, 2, \dots$,

$$\lambda P(Z+1=k) = k P(Z=k)$$

$$\text{i.e. } P(Z=k) = \frac{\lambda}{k} P(Z=k-1)$$

$$= \dots = \frac{\lambda^k}{k!} P(Z=0).$$

Since $\sum_{k=0}^{\infty} P(Z=k) = 1$, we have

$$1 = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} P(Z=0) = e^{\lambda} P(Z=0)$$

$$\Rightarrow P(Z=0) = e^{-\lambda}$$

$$\Rightarrow P(Z=k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k=0,1,2\ldots$$

$$\Rightarrow Z \sim Po(\lambda).$$

From Prop 3.1, $\mathcal{L}(W) \simeq Po(\lambda)$ if and only if

$$E\{\lambda f(W+1) - W f(W)\} \simeq 0.$$

<u>Theorem 3.2</u>

$X_1, \ldots X_n$ indep, $P(X_i=1) = 1 - P(X_i=0) = p_i$,

$W = \sum_{i=1}^{n} X_i$, $W^{(i)} \triangleq W - X_i$,

$\lambda = EW = \sum_{i=1}^{n} p_i$, $Z \sim Po(\lambda)$.

then $d(\mathcal{L}(W), Po(\lambda)) \le \frac{1-e^{-\lambda}}{\lambda} \sum_{i=1}^{n} p_i^2$

$$\le (1 \wedge \tfrac{1}{\lambda}) \sum_{i=1}^{n} p_i^2. \quad (3.2)$$

Proof.

Step 1 For bdd $f : \mathbb{Z}^+ \to \mathbb{R}$,

$$EWf(W) = \sum_{i=1}^{n} EX_i f(W) = \sum_{i=1}^{n} p_i E f(W^{(i)}+1)$$

$$= \sum_{i=1}^{n} p_i E[f(W^{(i)}+1) - f(W+1)] + \lambda E f(W+1)$$

$$= \sum_{i=1}^{n} p_i^2 E[f(W^{(i)}+1) - f(W^{(i)}+2)] + \lambda f(W+1)$$

$$\therefore E\{\lambda f(W+1) - Wf(W)\} = \sum_{i=1}^{n} p_i^2 E \Delta f(W^{(i)}+1) \quad (3.3)$$

where $\Delta f(x) = f(x+1) - f(x)$.

Step 2 Let $f = f_A$, a bounded solution of

$$\lambda f(w+1) - wf(w) = I_A(w) - P(Z \in A) \quad (3.4)$$

where $A \subset \mathbb{Z}^+$. Then

$$P(W \in A) - P(Z \in A) = \sum_{i=1}^{n} p_i^2 E \Delta f_A(W^{(i)}+1) \quad (3.5).$$

$$d(\mathcal{L}(W), P_0(\lambda)) = \sup_{A \subset \mathbb{Z}^+} \left| \sum_{i=1}^{n} p_i^2 E \Delta f_A(W^{(i)}+1) \right| \quad (3.6)$$

Step 3 Solving (3.4)

$$\lambda e^{-\lambda}\frac{\lambda^{k}}{k!}f(k+1)-\lambda e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}f(k)$$

$$=P(Z=k)\left[I_A(w)-P(Z\in A)\right]$$

where $\frac{\lambda^{k-1}}{(k-1)!}f(k)=0$ for $k=0$.

Sum over $k=0,1,\ldots,w-1$.

Then for $w\geq 1$,

$$f_A(w)=\frac{1}{\lambda P(Z=w-1)}E\left[I_A(Z)-P(Z\in A)\right]I(Z\leq w-1)$$

$$=-\frac{1}{\lambda P(Z=w-1)}E\left[I_A(Z)-P(Z\in A)\right]I(Z\geq w) \quad (3.7)$$

~~$f_A(w)$ is unique except at $w=0$.~~

But $f_A(0)$ does not enter into consideration.

By Prop. 3.3 below, for all $A\subset\mathbb{Z}^+$ & $w\geq 1$,

$$|\Delta f_A(w)|\leq\frac{1-e^{-\lambda}}{\lambda}\leq 1\wedge\lambda^{-1}.$$

Applying this to (3.6),

$$d(\mathcal{L}(W),P_0(\lambda))\leq\frac{1-e^{-\lambda}}{\lambda}\sum_{i=1}^{n}p_i^2\leq(1\wedge\lambda^{-1})\sum_{i=1}^{n}p_i^2$$

Proposition 3.3. For $A \subset \mathbb{Z}^+$ and $w \geq 1$,

(i) $|f_A(w)| \leq 1 \wedge 1.4\lambda^{-\frac{1}{2}}$

(ii) $|f_{\{0\}}(w)| = \frac{1-e^{-\lambda}}{\lambda} \leq 1 \wedge \lambda^{-1}$

(iii) $|\Delta f_A(w)| \leq \frac{1-e^{-\lambda}}{\lambda} \leq 1 \wedge \lambda^{-1}$

Remarks. (i) & (iii) due to Barbour & Eagleson (Adv. Appl. Prob (1983)). Now by a probabilistic approach $|f_A(w)| \leq 1 \wedge \lambda^{-\frac{1}{2}}$.

- (ii) due to Arratia, Goldstein & Gordon (Ann. Prob. (1989))

Proof.

(i) Following Chen (Ann Prob (1975)), we prove a weaker version: $\|f_A\| \leq 2(1 \wedge \lambda^{-\frac{1}{2}})$.

For $w \geq 1$,

$$f_A(w) = \frac{1}{\lambda P(Z=w-1)}\Big[P(Z \in A, Z \leq w-1) - P(Z \in A)P(Z \leq w-1)\Big]$$

$$= \frac{1}{\lambda P(Z=w-1)}\Big[P(Z \in A, Z \leq w-1)P(Z \geq w) - P(Z \in A, Z \geq w)P(Z \leq w-1)\Big]$$

So for $w \geq 1$

$$|f_A(w)| \leq \frac{P(Z \leq w-1)P(Z \geq w)}{\lambda P(Z=w-1)}$$

For $\lambda \geq w$ & $w \geq 2$,

$$|f_\lambda(w)| \leq \frac{P(Z \leq w-1)}{\lambda P(Z = w-1)}$$

$$= (w-1)!\,\lambda^{-w} \sum_{k=0}^{w-1} \frac{\lambda^k}{k!}$$

$$= \sum_{k=0}^{w-1} \frac{(w-1)\cdots(k+1)}{\lambda^{w-k}}$$

$$= \sum_{\ell=0}^{w-1} \frac{(w-1)\cdots(w-\ell)}{\lambda^{\ell+1}} \quad \text{where } \ell = w-1-k$$

$$\leq \sum_{\ell=0}^{[\lambda]-1} \frac{(\lambda-1)\cdots(\lambda-\ell)}{\lambda^{\ell+1}}$$

$$= \left(\sum_{\ell=0}^{[\varphi(\lambda)]-1} + \sum_{\ell=[\varphi(\lambda)]}^{[\lambda]-1} \right) \frac{(\lambda-1)\cdots(\lambda-\ell)}{\lambda^{\ell+1}}$$

where $1 \leq [\varphi(\lambda)] \leq [\lambda]$

$$\leq \frac{[\varphi(\lambda)]}{\lambda} + \sum_{m=0}^{\infty} \frac{(\lambda-1)\cdots(\lambda-[\varphi(\lambda)]-m)}{\lambda^{[\varphi(\lambda)]+m+1}}$$

where $m = \ell - [\varphi(\lambda)]$

$$\leq \frac{[\varphi(\lambda)]}{\lambda} + \sum_{m=0}^{\infty} \frac{(\lambda-\varphi(\lambda))^m}{\lambda^{m+1}}$$

$$\leq \frac{\varphi(\lambda)}{\lambda} + \frac{1}{\varphi(\lambda)}$$

$$= \frac{2}{\lambda^{\frac{1}{2}}} \quad \text{by taking } \varphi(\lambda) = \lambda^{\frac{1}{2}}$$

$$\leq 2(1 \wedge \lambda^{-\frac{1}{2}}).$$

For $\lambda \geq w$ & $w = 1$,

$$|f_A(w)| \leq \frac{1}{\lambda} \leq 2(1 \wedge \lambda^{-\frac{1}{2}})$$

For $0 < \lambda \leq w$ & $w \geq 2$,

$$|f_A(w)| \leq \frac{P(Z \geq w)}{\lambda P(Z = w-1)}$$

$$= (w-1)!\, \lambda^{-w} \sum_{k=w}^{\infty} \frac{\lambda^k}{k!}$$

$$= \sum_{k=w}^{\infty} \frac{\lambda^{k-w}}{w(w+1)\cdots k}$$

$$= \sum_{l=0}^{\infty} \frac{\lambda^l}{w(w+1)\cdots(w+l)} \quad \text{where } l = k-w$$

$$\leq \sum_{l=0}^{\infty} \frac{w^{l-1}}{(w+1)\cdots(w+l)}$$

$$= \left(\sum_{l=0}^{[w^{\frac{1}{2}}]-2} + \sum_{l=[w^{\frac{1}{2}}]-1}^{\infty} \right) \frac{w^{l-1}}{(w+1)\cdots(w+l)}$$

No

Date

$$\leq \frac{1}{\omega}\left\{[\omega^{\frac{1}{2}}]-1+1+\sum_{l=[\omega^{\frac{1}{2}}]}^{\infty}\frac{\omega^{l}}{(\omega+1)\cdots(\omega+l)}\right\}$$

$$=\frac{1}{\omega}\left\{[\omega^{\frac{1}{2}}]+\sum_{m=0}^{\infty}\frac{\omega^{[\omega^{\frac{1}{2}}]+m}}{(\omega+1)\cdots(\omega+[\omega^{\frac{1}{2}}]+m)}\right\}$$

where $m=l-[\omega^{\frac{1}{2}}]$

$$=\frac{1}{\omega}\left\{[\omega^{\frac{1}{2}}]+\frac{\omega^{[\omega^{\frac{1}{2}}]}}{(\omega+1)\cdots(\omega+[\omega^{\frac{1}{2}}])}\sum_{m=0}^{\infty}\frac{\omega^{m}}{(\omega+[\omega^{\frac{1}{2}}]+1)\cdots(\omega+[\omega^{\frac{1}{2}}]+m)}\right\}$$

$$\leq \frac{1}{\omega}\left\{[\omega^{\frac{1}{2}}]+\frac{\omega}{\omega+[\omega^{\frac{1}{2}}]}\sum_{m=0}^{\infty}\frac{\omega^{m}}{(\omega+\omega^{\frac{1}{2}})^{m}}\right\}$$

Since $\omega^{\frac{1}{2}}\leq[\omega^{\frac{1}{2}}]+1$

$$=\frac{1}{\omega}\left\{[\omega^{\frac{1}{2}}]+\frac{\omega}{\omega+[\omega^{\frac{1}{2}}]}\cdot\frac{\omega+\omega^{\frac{1}{2}}}{\omega^{\frac{1}{2}}}\right\}$$

$$=\frac{1}{\omega^{\frac{1}{2}}}\left\{[\omega^{\frac{1}{2}}]\omega^{-\frac{1}{2}}+\frac{\omega+\omega^{\frac{1}{2}}}{\omega+[\omega^{\frac{1}{2}}]}\right\}$$

$$=\frac{1}{\omega^{\frac{1}{2}}}\left\{2+\frac{[\omega^{\frac{1}{2}}]-\omega^{\frac{1}{2}}}{\omega^{\frac{1}{2}}}+\frac{\omega^{\frac{1}{2}}-[\omega^{\frac{1}{2}}]}{\omega+[\omega^{\frac{1}{2}}]}\right\}$$

$$\leq\frac{2}{\omega^{\frac{1}{2}}}\leq 2(1\wedge\lambda^{-\frac{1}{2}})$$

For $0<\lambda\le w$ & $w\ge 1$,

$$|f_A(w)| = \frac{P(Z\ge 1)}{\lambda} = \frac{1-P(Z=0)}{\lambda}$$

$$= \frac{1-e^{-\lambda}}{\lambda} \le (1\wedge\lambda^{-1}) \le 2(1\wedge\lambda^{-\frac{1}{2}})$$

Note: the upper bound $\dfrac{P(Z\le w-1)P(Z\ge w)}{\lambda P(Z=w-1)}$

on $|f_A(w)|$ is attained when $A=\{0,1,\dots,w-1\}$

(ii) For $w\ge 1$,

$$f_{\{0\}}(w) = \frac{P(Z=0)-P(Z=0)P(Z\le w-1)}{\lambda P(Z=w-1)}$$

Now $P(Z\le w-1) = P(N_\lambda(1)\le w-1)$

$$= P(T_w>1) = \int_1^\infty \frac{\lambda^w x^{w-1}}{(w-1)!}e^{-\lambda x}dx$$

$$= \int_\lambda^\infty \frac{x^{w-1}}{(w-1)!}e^{-x}dx$$

Alternatively,

$$\frac{d}{d\lambda}P(Z\le w-1) = -P(Z=w-1) = -e^{-\lambda}\frac{\lambda^{w-1}}{(w-1)!}$$

No

Date

18

$$\therefore \; P(Z \le w-1) = 1 - \int_0^{\lambda} e^{-x} \frac{x^{w-1}}{(w-1)!} dx$$

using the boundary condition that

$P(Z \le w-1) = 1$ if $\lambda = 0$.

$\therefore$ For $w \ge 1$,

$$f_{\{0\}}(w) = \frac{P(Z=0)}{\lambda P(Z=w-1)} \left[\int_0^{\lambda} e^{-x} \frac{x^{w-1}}{(w-1)!} dx \right]$$

$$= \frac{1}{\lambda} \int_0^{\lambda} \left(\frac{x}{\lambda}\right)^{w-1} e^{-x} dx$$

$$\therefore \; \frac{1-e^{-\lambda}}{\lambda} = f_{\{0\}}(1) \ge f_{\{0\}}(2) \ge \cdots$$

(iii) Write $f_{\{i\}}$ as f_i.

From (3.7.), for $j \ge 0$ and $w \ge 1$,

$$f_j(w) = \begin{cases} -\dfrac{P(Z=j)\,P(Z \le w-1)}{\lambda P(Z=w-1)} & \text{if } w \le j, \\[2ex] \dfrac{P(Z=j)\,P(Z \ge w)}{\lambda P(Z=w-1)} & \text{if } w \ge j+1 \end{cases}$$

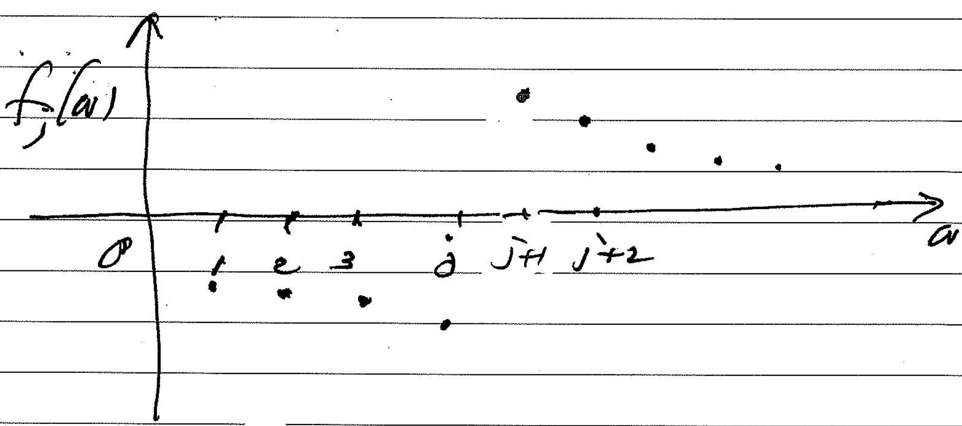
Claim (will prove later):

For $j \geq 0$,

$$f_j(j+1) > f_j(j+2) > \cdots > 0$$

For $j \geq 1$,

$$0 > f_j(1) > f_j(2) > \cdots > f_j(j).$$



Hence for $w \geq 1$,

$$\Delta f_j(w) = f_j(w+1) - f_j(w) \begin{cases} > 0 & \text{if } w = j \\ < 0 & \text{if } w \neq j \end{cases}.$$

Now for $A \subset \mathbb{Z}^+$ & $w \geq 1$

$$f_A(w) = \sum_{j \in A} f_j(w)$$

So $\Delta f_A(w) = \sum_{j \in A} \Delta f_j(w) \leq \Delta f_w(w).$

No

Date

Lekeàrzè

$$\Delta f_{A^c}(w) \le \Delta f_w(w)$$

But $\Delta f_A(w) + \Delta f_{A^c}(w) = \Delta f_{\mathbb{Z}^+}(w) = 1$

So $-\Delta f_w(w) \le -\Delta f_{A^c}(w) = \Delta f_A(w) \le \Delta f_w(w)$.

Hence $|\Delta f_A(w)| \le \Delta f_w(w)$.

Now $\Delta f_w(w) = f_w(w+1) - f_w(w)$

$$= \frac{P(Z=w)P(Z\ge w+1)}{\lambda P(Z=w)} + \frac{P(Z=w)P(Z\le w-1)}{\lambda P(Z=w-1)}$$

$$= \frac{1}{\lambda}\left\{P(Z\ge w+1) + \frac{\lambda}{w}P(Z\le w-1)\right\}$$

$$= \frac{1}{\lambda}\left\{P(Z\ge w+1) + \frac{1}{w}\lambda E I(Z+1\le w)\right\}$$

$$= \frac{1}{\lambda}\left\{P(Z\ge w+1) + \frac{1}{w}E Z I(Z\le w)\right\}$$

$$\le \frac{1}{\lambda}\left\{P(Z\ge w+1) + P(1\le Z\le w)\right\}$$

$$= \frac{1}{\lambda}P(Z\ge 1) = \frac{1-e^{-\lambda}}{\lambda} \le 1\wedge\lambda^{-1}.$$

No
Date

Hence $|\Delta f_A(w)| \le \frac{1-e^{-\lambda}}{\lambda} \le (\lambda)^{-1}$.

Proof of claim:

For $w \ge j+1$,

$$\frac{f_j(w)}{f_{j+1}(w+1)} = \frac{P(Z \ge w)}{P(Z = w-1)} \cdot \frac{P(Z = w)}{P(Z \ge w+1)}$$

$$= \frac{\lambda}{w} \cdot \frac{P(Z \ge w)}{P(Z \ge w+1)} = \frac{\lambda E I(Z+1 \ge w+1)}{w P(Z \ge w+1)}$$

$$= \frac{E Z I(Z \ge w+1)}{w P(Z \ge w+1)} > 1.$$

For $1 \le w \le j$,

$$\frac{f_j(w)}{f_{j+1}(w+1)} = \frac{P(Z \le w-1)}{P(Z = w-1)} \cdot \frac{P(Z = w)}{P(Z \le w)}$$

$$= \frac{\lambda}{w} \cdot \frac{P(Z \le w-1)}{P(Z \le w)} = \frac{\lambda E I(Z+1 \le w)}{w P(Z \le w)}$$

$$= \frac{E Z I(Z \le w)}{w P(Z \le w)} < 1.$$

§4 Dependent events

4.1 Local approach

Stein (6th Berkeley Symposium (1972))

Chen (Ann. Prob. (1975))

Arratia, Goldstein & Gordon (Ann. Prob (1989), Stat Science (1990)).

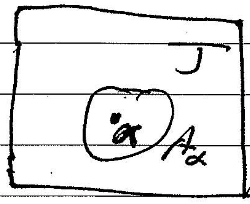
Theorem 4.1

$\{X_\alpha : \alpha \in J\}$ random indicators, not necessarily indep.

$P(X_\alpha = 1) = 1 - P(X_\alpha = 0) = p_\alpha$

$W = \sum_{\alpha \in J} X_\alpha$, $\lambda = EW = \sum_{\alpha \in J} p_\alpha$, $Z \sim Po(\lambda)$

$\forall \alpha \in J$, let $A_\alpha \subsetneq J$ s.t. $\alpha \in A_\alpha$



Then

$$d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda^{-1})(b_1 + b_2) + (1 \wedge 1.4\lambda^{-\frac{1}{2}})\, b_3 \quad (4.1)$$

& $$|P(W=0) - e^{-\lambda}| \le (1 \wedge \lambda^{-1})(b_1 + b_2 + b_3) \quad (4.2)$$

where $b_1 = \sum_{\alpha \in J} \sum_{\beta \in A_\alpha} p_\alpha p_\beta$, $b_2 = \sum_{\alpha \in J} \sum_{\alpha \neq \beta \in A_\alpha} E X_\alpha X_\beta$,

$$b_3 = \sum_{\alpha \in J} E \left| E\left(X_\alpha - p_\alpha \,\Big|\, \sum_{\beta \in A_\alpha^c} X_\beta\right) \right|$$

Proof. Let $W^{(\alpha)} = W - X_\alpha \in V_\alpha = \sum_{\beta \in A_\alpha^c} X_\beta$.

$\forall$ bdd $f : \mathbb{Z}^+ \to \mathbb{R}$

$$E Wf(W) = \sum_{\alpha \in J} E X_\alpha f(W^{(\alpha)}+1)$$

$$= \sum_{\alpha \in J} E X_\alpha [f(W^{(\alpha)}+1) - f(V_\alpha+1)]$$

$$+ \sum_{\alpha \in J} E(X_\alpha - p_\alpha) f(V_\alpha+1)$$

$$+ \sum_{\alpha \in J} p_\alpha E[f(V_\alpha+1) - f(W+1)]$$

$$+ \lambda E f(W+1)$$

So

$$E\{\lambda f(W+1) - W f(W)\}$$

$$= \sum_{\alpha \in J} p_\alpha E[f(V_\alpha + \sum_{\alpha \in A_\alpha} X_\beta + 1) - f(V_\alpha+1)]$$

$$- \sum_{\alpha \in J} E X_\alpha [f(V_\alpha + \sum_{\alpha \neq \beta \in A_\alpha} X_\beta + 1) - f(V_\alpha+1)]$$

$$- \sum_{\alpha \in J} E[f(V_\alpha+1) E(X_\alpha - p_\alpha \mid V_\alpha)]$$

Now let $f = f_A$, a bounded solution of

$$\lambda f(w+1) - w f(w) = I_A(w) - P(Z \in A), \quad A \subset \mathbb{Z}^+.$$

~~then~~ $P(W \in A) - P(Z \in A)$

$$= \sum_{\alpha \in J} p_i \, E\Big[f_A\Big(V_\alpha + \sum_{\beta \in A_\alpha} X_\beta + 1\Big) - f_A(V_\alpha + 1)\Big]$$

$$- \sum_{\alpha \in J} EX_\alpha \, E\Big[f_A\Big(V_\alpha + \sum_{\alpha \neq \beta \in A_\alpha} X_\beta + 1\Big) - f_A(V_\alpha + 1)\Big]$$

$$- \sum_{\alpha \in J} E\big[f_A(V_\alpha + 1) E(X_\alpha - p_\alpha \mid V_\alpha)\big] \quad (4.3)$$

Observe that for any finite $B \subset \mathbb{Z}^+$ & any r.v. $U \geq 1$

$$\Big| f_A\Big(U + \sum_{\beta \in B} X_\beta\Big) - f_A(U) \Big|$$

$$= \Big| f_A\Big(U + \sum_{\beta=1}^{k} X_\beta\Big) - f_A(U) \Big| \quad \Big(\text{by ordering } \sum_{\beta \in B} X_\beta \text{ as } \sum_{\beta=1}^{k} X_\beta\Big)$$

$$\leq \sum_{i=1}^{k} \Big| f_A\Big(U + \sum_{\beta=1}^{i} X_\beta\Big) - f_A\Big(U + \sum_{\beta=1}^{i-1} X_\beta\Big) \Big|$$

$$= \sum_{i=1}^{k} \Big| X_i \Big[f_A\Big(U + \sum_{\beta=1}^{i-1} X_\beta + 1\Big) - f_A\Big(U + \sum_{\beta=1}^{i-1} X_\beta\Big)\Big] \Big|$$

$$\leq \Big(\sum_{i=1}^{k} X_i\Big) \sup_{w \geq 1} |\Delta f_A(w)|$$

$$= (1 \wedge \lambda^{-1}) \sum_{\beta \in B} X_\beta \quad (\text{by Prop } 3.3 \text{ (iii)}) \quad (4.4)$$

Applying (4.4) and Prop 3.3 (ii) to (4.3),

$$d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda^{-1}) \sum_{\alpha \in J} \mu_\alpha E\Big(\sum_{\beta \in A_\alpha} X_\beta\Big)$$

$$+ (1 \wedge \lambda^{-1}) \sum_{\alpha \in J} E X_\alpha \Big(\sum_{\alpha \neq \beta \in A_\alpha} X_\beta\Big)$$

$$+ (1 \wedge (1.4\lambda)^{-\frac{1}{2}}) \sum_{\alpha \in J} E\left|E(X_\alpha - \mu_\alpha \mid V_\alpha)\right|$$

$$= (1 \wedge \lambda^{-1})(b_1 + b_2) + (1 \wedge 1.4\lambda^{-\frac{1}{2}})\, b_3$$

This proves (4.1). For (4.2), let $A = \{0\}$, then apply (4.4) and Prop 3.3 (ii) to (4.3).

<u>Definition</u>:
$\{X_\alpha : \alpha \in J\}$ is locally dependent if $\forall \alpha \in J$ $\exists A_\alpha \subsetneq J$ s.t. $\alpha \in A_\alpha$ and X_α is indep of $\{X_\beta : \beta \in A_\alpha^c\}$.

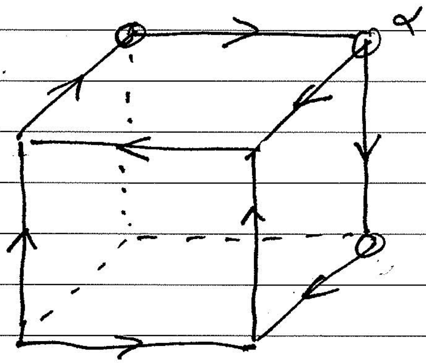
A_α is called a dependent ~~set~~ neighbourhood of α & $\{X_\beta : \beta \in A_\alpha\}$ is dependent set of X_α.

<u>Note</u>: Under local dependence with dependent set A_α, $b_3 = 0$.

Applications

(1) A random graph problem

n-dim cube $\{0,1\}^n$, 2^n vertices, $n2^{n-1}$ edges.



Each edge is assigned a random direction by tossing a fair coin.

Let $W = W(k,n) =$ no. of vertices at which exactly k edges point outward.

Let J = set of all 2^n vertices,

$X_\alpha = I($vertex α has exactly k of its edges directed outward$)$

Then $W = \sum_{\alpha \in J} X_\alpha$,

$\{X_\alpha : \alpha \in J\}$ locally dependent

$A_\alpha = \{\beta : |\beta - \alpha| \le 1\}$ dependent set

where $|\beta - \alpha|$ is the distance between

$\beta = (\beta_1, \dots, \beta_n)$ and $\alpha = (\alpha_1, \dots, \alpha_n)$

and is defined by

$$|\beta - \alpha| = |\alpha - \beta| = \sum_{i=1}^{n} |\alpha_i - \beta_i| ,$$

$$P(X_\alpha = 1) = 1 - P(X_\alpha = 0) = p_\alpha = \frac{\binom{n}{k}}{2^n}$$

$$\lambda = EW = 2^n \cdot \frac{\binom{n}{k}}{2^n} = \binom{n}{k}$$

By local dependence $b_3 = 0$.

For $|\alpha - \beta| = 1$,

$$EX_\alpha X_\beta = E(X_\alpha X_\beta / \alpha \to \beta)\frac{1}{2} + E(X_\alpha X_\beta / \alpha \leftarrow \beta)\frac{1}{2}$$

$$= \frac{\binom{n-1}{k}}{2^{n-1}} \cdot \frac{\binom{n-1}{k-1}}{2^{n-1}} \cdot \frac{1}{2} + \text{same}$$

$$= \frac{\binom{n-1}{k}\binom{n-1}{k}}{2^{n-2}} = \frac{\binom{n}{k}^2}{2^{n-2}} \cdot \frac{k}{n}\left(1 - \frac{k}{n}\right)$$

$$\leq \frac{\binom{n}{k}^2}{2^{2n}}$$

So $b_2 \leq 2^n \cdot n \cdot \frac{\binom{n}{k}^2}{2^{2n}} = \frac{n\binom{n}{k}^2}{2^n}$.

Finally

$$b_1 = 2^n (n+1) \cdot \frac{\binom{n}{k}^2}{2^{2n}} = \frac{(n+1)\binom{n}{k}^2}{2^n}$$

Hence $d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda^{-1}) \cdot \dfrac{(2n+1)\binom{n}{k}^2}{2^n}$,

$$\le \frac{1}{\binom{n}{k}} \, \frac{(2n+1)\binom{n}{k}^2}{2^n} = \frac{(2n+1)\binom{n}{k}}{2^n} \to 0$$

If $1 \le k \le n-1$, then $\lambda = \binom{n}{k} \to \infty$.

In this case we have the following central limit theorem

$$P\left(\frac{W - \binom{n}{k}}{\sqrt{\binom{n}{k}}} \le x \right) \to \Phi(x).$$

(2) The birthday problem

n balls (people) are uniformly and independently distributed into d boxes (days of the year). $(n<d)$.

P(At least one box contains k or more balls) $= ?$

For $k=2$,

$$\text{prob.} = 1 - \frac{d(d-1)\cdots(d-n+1)}{d^n}$$

For $k=3$,

$$\text{prob} = 1 - \frac{1}{d^n} \sum_{i+2j=n} \frac{d!}{i!\,j!\,(d-i-j)!} \cdot \frac{n!}{2^j}.$$

For $k \geq 4$, don't know.

Let $J = \{\alpha \subset \{1, 2, \ldots, n\} : |\alpha| = k\}$

$X_\alpha = I$(the balls indexed by α all go into same box)

Then $\{X_\alpha : \alpha \in J\}$ is locally dependent with

$$A_\alpha = \{\beta \in J : \alpha \cap \beta \neq \phi\}$$

$$p_\alpha = \frac{1}{d^{k-1}} \quad \& \quad \lambda = |J| p_\alpha = \frac{\binom{n}{k}}{d^{k-1}}$$

Let $W = \sum_{\alpha \in J} X_\alpha$.

Then $\{$No box gets k or more balls$\} = \{W = 0\}$

Take $n, d \to \infty$ s.t. $\lambda \simeq 1$.

Then $|P$(At least one box gets k or more balls)
$- (1 - e^{-\lambda})|$

$= |P(W=0) - e^{-\lambda}| \leq (1 \wedge \lambda^{-1})(b_1 + b_2 + b_3)$

(by (4.2)).

Now local dependence with A_α as dependent sets
$\Rightarrow b_3 = 0$.

From the definition of A_α, $|A_\alpha| = \binom{n}{k} - \binom{n-k}{k}$

So $b_1 = p_\alpha^2 |\Im||A_\alpha| = \dfrac{\lambda^2|A_\alpha|}{|\Im|} = \dfrac{\lambda^2\left[\binom{n}{k}-\binom{n-k}{k}\right]}{\binom{n}{k}}$

$$= \lambda^2\left[1-\left(1-\frac{k}{n}\right)\left(1-\frac{k}{n-1}\right)\cdots\left(1-\frac{k}{n-k+1}\right)\right]$$

$$\le \frac{\lambda^2 k^2}{n}.$$

For the case $k=2$, X_α & X_β are pairwise independent. (another example that pairwise indep $\not\Rightarrow$ indep.)

$$\left(EX_\alpha X_\beta = E(X_\alpha \mid X_\beta = 1)\frac{1}{d^{k-1}} = \frac{1}{d}\frac{1}{d} = p_\alpha p_\beta .\right)$$

So $b_2 = |\Im|(|A_\alpha|-1)\, EX_\alpha X_\beta$

$$\le |\Im||A_\alpha|\, p_\alpha^2 = b_1 .$$

Hence

$$\left|P(\text{at least one box gets 2 or more balls}) - (1-e^{-\lambda})\right| \le 2(1\wedge\lambda^{-1})\, b_1$$

$$\le \frac{2\lambda^2(1\wedge\lambda^{-1})}{n} = O\left(\frac{1}{n}\right)$$

For general k,

$$b_2 = \sum_{j=1}^{k-1}\binom{n}{k}\binom{k}{j}\binom{n-k}{k-j} d^{1+j-2k}$$

where the j^{th} term is the contribution to b_2 from pairs (α,β) with $|\alpha\cap\beta| = j$ and $E X_\alpha X_\beta = d^{1+j-2k}$.

The dominant contribution to b_2 comes from the term with $j = k-1$.

Hence $b_2 = O\left(\frac{n^{k+1}}{d^k}\right)$.

Now $\lambda \simeq 1 \Rightarrow \frac{n^k}{d^{k-1}} \simeq 1$

$$\Rightarrow \frac{n^{k+1}}{d^k} \simeq \frac{n}{d} \simeq n^{-\frac{1}{k-1}}$$

~~Hence~~

$$|P(\text{At least one box gets } k \text{ or more balls}) - (1-e^{-\lambda})| = O\left(n^{-\frac{1}{k-1}}\right).$$

(3) The length of the longest head run

Toss a coin repeatedly

$P(\text{Head}) = p \quad (0<p<1)$

Let R_n = length of the longest run of heads starting from within the 1st n tosses.

What is the asymptotics of R_n as $n \to \infty$?

Let $Z_1, Z_2, \dots$ be i.i.d. with

$$P(\underset{\text{Head}}{\underbrace{Z_i = 1}}) = p = 1 - P(\underset{\text{Tail}}{\underbrace{Z_i = 0}})$$

Let $J = \{1, 2, \dots, n\}$ and let $t \geq 1$.

Define $Y_i = Z_i Z_{i+1} \cdots Z_{i+t-1}$ for $i = 1, 2, \dots, n$

$$Z_{i-2}, Z_{i-1}, \underbrace{Z_i, Z_{i+1}, \dots, Z_{i+t-1}}_{Y_i}, Z_{i+t}, Z_{i+t+1}$$

Define

$$X_i = \begin{cases} Y_1 & \text{if } i = 1 \\ (1 - Z_{i-1}) Y_i & \text{if } i \geq 2 \end{cases}$$

Then

$$P(X_i = 1) = 1 - P(X_i = 0) = \begin{cases} p^t & \text{if } i = 1 \\ p^t(1-p) & \text{if } i \geq 2 \end{cases}$$

Define $W = \sum_{i=1}^{n} X_i$.

Then $\{R_n < t\} = \{W = 0\}$

Also $\lambda = EW = p^t\{(n-1)(1-p)+1\}$

& $Z \sim Po(\lambda)$.

Define $A_i = \{j \in J : |i-j| \le t\}$, $i = 1, \dots, n$.

Then $\{X_i : i \in J\}$ is locally dependent

with ~~dependent~~ neighbourhoods $\{A_i : i \in J\}$.

So $b_3 = 0$.

For b_2, observe that $EX_iX_j = 0$ for $|i-j| \le t$,

and so $b_2 = 0$.

$$b_1 = \sum_{i=1}^{n} \sum_{j \in A_i} p_i p_j$$

$$= \sum_{i=1}^{n} \sum_{\substack{j \in A_i \\ j \ne 1}} p_i p_j + \sum_{i=1}^{n} \sum_{\substack{j \in A_i \\ j = 1}} p_i p_j$$

$$= \sum_{i=1}^{n} \sum_{\substack{j \in A_i \\ j \ne 1}} p_i p^t(1-p) + \sum_{i=1}^{n} \sum_{\substack{j \in A_i \\ j = 1}} p_i p^t$$

$$\le \lambda(2t+1)p^t(1-p) + \lambda p^t$$

$$< \frac{\lambda^2(2t+1)}{n} + \lambda p^t$$

Hence $|P(R_n < t) - e^{-\lambda}|$

$$= |P(W=0) - e^{-\lambda}| \leq (1 \wedge \lambda^{-1}) b_1$$

$$\leq (1 \wedge \lambda^{-1})\left(\frac{\lambda^2(2t+1)}{n} + \lambda p^t\right)$$

As $n \to \infty$ want λ to be bounded away from 0 and from ∞ and at the same time the upper bound $\to 0$.

This is equivalent to $\log \lambda$ being bounded

i.e. $t - \log_{\frac{1}{p}}(n(1-p))$ bounded

which is equivalent to $t - [\log_{\frac{1}{p}}(n(1-p))]$ being bounded.

Let $t = [\log_{\frac{1}{p}}(n(1-p))] + c$

where c is a fixed integer.

Then $|P(R_n - [\log_{\frac{1}{p}}(n(1-p))] < c) - e^{-\lambda}|$

$\to 0.$

No

Date

But $\lambda = \rho^{c + \log_{\rho}((n-1)(1-p)+1)}$

$$= \rho^{c - \log_{\frac{1}{\rho}}((n-1)(1-p)+1)}$$

$$= \rho^{c + [\log_{\frac{1}{\rho}}(n(1-p))] - \log_{\frac{1}{\rho}}(n(1-p)) + \underbrace{\log_{\frac{1}{\rho}}(n(1-p)) - \log_{\frac{1}{\rho}}((n-1)(1-p)+1)}_{= \log_{\frac{1}{\rho}} \frac{n(1-p)}{(n-1)(1-p)+1} \to 0 \text{ as } n\to\infty}}$$

Hence $P(R_n - [\log_{\frac{1}{\rho}}(n(1-p))] < c)$

$$\to e^{-\rho^{c-r}} \text{ along a subsequence}$$

~~if and only if~~

$\log_{\frac{1}{\rho}}(n(1-p)) - [\log_{\frac{1}{\rho}}(n(1-p))] \to r \in [0,1]$

along the same subsequence.

No limit but we have an approximation result.

Recall that under independence

$$d(\mathcal{L}(W), \mathcal{L}(Z)) \leq (1 \wedge \lambda^{-1}) \sum_{i=1}^{n} p_i^2$$

$$= (1 \wedge \lambda) \left\{ 1 - \frac{Var(W)}{\lambda} \right\}$$

$$= (1 \wedge \lambda) \left\{ \frac{Var(W)}{\lambda} - 1 + \frac{2}{\lambda} \sum_{i=1}^{n} p_i^2 \right\} \quad (4.5)$$

Under local dependence

$$d(\mathcal{L}(W), \mathcal{L}(Z)) \leq (1 \wedge \lambda) \left\{ \frac{Var(W)}{\lambda} - 1 + \frac{2}{\lambda} \sum_{\alpha \in J} \sum_{\beta \in A_\alpha} p_\alpha p_\beta \right\} \quad (4.6)$$

No

Date

(4) Palindromes in DNA

Palindrome: A word or a phrase that is the same whether you read it backwards or forwards.

Example: LEVEL

REFER

ABLE WAS I ERE I SAW ELBA

Elba – A mountainous island off the W. of Italy, in the Mediterranean. Napoleon Bonaparte's first place of exile (1814–15).

DNA: A string of letters taken from the alphabet {A, T, C, G}

Most DNA are double-stranded

A, T & C, G are complementary pairs.

C G T C A T G G A T C C A G T A G
G C A G T A C C T A G G T C A T C

base pairs

Human genome: 3 billion base pairs

Bacteria genome: 3 million base pairs

Virus genome: 200,000 base pairs.

A palindrome on DNA is a word which reads exactly the same as its reverse complementary sequence.

C G T C A T G G A T C C A G T A G

G C A G T A C C T A G G T C A T C

Centre

palindrome of length 8

palindrome length must be even

Assume the letters on a DNA are random & independent.

$p_A = p_T$ & $p_C = p_G$

$p_A + p_T + p_C + p_G = 2(p_A + p_C) = 1.$

$P(\exists$ a palindrome of length $\geq 2L$ with centre $i) = \theta^L$

where $\theta = 2(p_A p_T + p_C p_G)$

$= 2(p_A^2 + p_C^2)$

For $L=5$ (length $=10$)

$p_A = p_T =$ & $p_C = p_G =$

θ^L

Palindromes are involved in a variety of biological processes.

(i) Recognition sites for bacterial restriction enzymes to cut foreign DNA

(ii) gene regulation

(iii) DNA replication.

Ref: Leung, Choi, Kao & Chen (2002):
Nonrandom clusters of palindromes in herpesvirus genome (Preprint)

A simplified problem

M base pairs

$n = M - 2L + 1$ possible centres of palindrome of length $\geq 2L$.

$J = \{1, 2, \ldots, n\}$

No
Date

$X_i = 1$ if i is a centre of palindrome $\geq 2L$
$= 0$ if not.

$P(X_i = 1) = 1 - P(X_i = 0) = \theta^L$

$\theta = 2(p_A^2 + p_C^2)$

$W = \sum_{i=1}^{n} X_i$ is the number of palindromes of lengths $\geq 2L$

$\{X_i : i \in J\}$ is locally dependent with dependent neighbourhoods

$A_i = \{1 \leq j \leq n : |j - i| < 2L\}$.

$\lambda = EW = n\theta^L$, $Z \sim Po(\lambda)$

$d(\mathcal{L}(W), \mathcal{L}(Z)) \leq (1 \wedge \lambda^{-1})(b_1 + b_2)$

$b_1 = \sum_{i \in J} \sum_{j \in A_i} p_i p_j = n(4L-1)\theta^{2L}$

$b_2 = \sum_{i \in J} \sum_{i \neq j \in A_i} E X_i X_j$

$\leq \sum_{i \in J} \sum_{i \neq j \in A_i} \theta^{\frac{3L}{2}}$ (Lemma 1 in paper by Leung, Choi, Xia & Chen)

$\leq 4nL\theta^{\frac{3L}{2}}$

So $d(\mathcal{L}(W), \mathcal{L}(Z)) \le (c_n \lambda^{-1})\, \delta\, n\, L\, \theta^{\frac{3L}{2}}$

want $\lambda \simeq 1$, as $n \to \infty$.

~~So $\log \lambda$ is bounded~~

Let $\log \lambda = c$

then $L = \log_{\frac{1}{\theta}} n - c$.

Usually $\lambda > 1$.

So $d(\mathcal{L}(W), \mathcal{L}(Z)) \le \dfrac{\delta\, n\, L\, \theta^{\frac{3L}{2}}}{\lambda}$

$= \delta L \theta^{\frac{L}{2}} \to 0$

4.2 Coupling approach

Barbour and Holst (Adv. Appl. Prob (1989))

Barbour, Holst & Janson, Poisson Approximation, 1992

Theorem 4.2

Let $\{X_\alpha : \alpha \in J\}$ be random indicators

with $P(X_\alpha = 1) = 1 - P(X_\alpha = 0) = p_\alpha$.

Suppose $\forall \alpha \in J$, $\exists \{Y_{\beta\alpha} : \beta \in J\}$ defined

on the same probability space as

$\{X_\alpha : \alpha \in J\}$ s.t.

$$\mathcal{L}(Y_{\beta\alpha} : \beta \in J) = \mathcal{L}(X_\beta : \beta \in J \mid X_\alpha = 1)$$

and $Y_{\beta\alpha} \le X_\beta$ for $\beta \in J$ (NR)

(respectively $Y_{\beta\alpha} \ge X_\beta$ for $\beta \in J$) (PR)

Let $W = \sum_{\alpha \in J} X_\alpha$, $\lambda = EW = \sum_{\alpha \in J} p_\alpha$,

& $Z \sim Po(\lambda)$.

Then $d(\mathcal{L}(W), \mathcal{L}(Z)) \le C(\lambda) \left\{ 1 - \frac{Var(W)}{\lambda} \right\}$ (4.7)

(respectively $d(\mathcal{L}(W), \mathcal{L}(Z)) \le C(\lambda) \left\{ \frac{Var(W)}{\lambda} - 1 + \frac{2}{\lambda} \sum_{\alpha \in J} p_\alpha^2 \right\}$)

(4.8)

Definitions

(i) $\{X_\alpha : \alpha \in J\}$ are negatively related if they satisfy NR

(ii) $\{X_\alpha : \alpha \in J\}$ are positively related if they satisfies PR.

Remark

If $\{X_\alpha : \alpha \in J\}$ are independent,

take $Y_{\beta\alpha} = \begin{cases} 1 & \text{if } \beta = \alpha \\ X_\beta & \text{if } \beta \neq \alpha. \end{cases}$

Hence independence $\Longrightarrow$ NR & PR (see (4.5) & (4.6))

Proof of Theorem 4.2 Let $V_\alpha = \sum_{\alpha \neq \beta \in J} Y_{\beta\alpha}$.

$$EWf(W) = \sum_{\alpha \in J} EX_\alpha f(W) = \sum_{\alpha \in J} p_\alpha E(f(W^{(\alpha)}+1) \mid X_\alpha = 1)$$

$$= \sum_{\alpha \in J} p_\alpha E f(V_\alpha + 1)$$

$$= \sum_{\alpha \in J} p_\alpha E[f(V_\alpha+1) - f(W+1)] + \lambda E f(W+1)$$

So

$$P(W \in A) - P(Z \in A) = E\{\lambda f_A(W+1) - W f_A(W)\}$$

$$= E \sum_{\alpha \in J} p_\alpha E[f_A(W+1) - f_A(V_\alpha+1)]$$

Under NR,

$$\left|\sum_{\alpha\in J} p_\alpha E\left[f_A(W+1) - f_A(V_\alpha+1)\right]\right|$$

$$\leq \sum_{\alpha\in J} (1\wedge\lambda^{-1})\, p_\alpha\, |W - V_\alpha|$$

$$= \sum_{\alpha\in J} (1\wedge\lambda^{-1})\, p_\alpha\, E\left[(W+1) - (V_\alpha+1)\right]$$

(Since $Y_{\beta\alpha} \leq X_\beta$)

$$= (1\wedge\lambda^{-1}) \left\{ \lambda^2 + \lambda - \sum_{\alpha\in J} p_\alpha E(W^{(\alpha)}+1 \mid X_\alpha = 1) \right\}$$

$$= (1\wedge\lambda^{-1}) \left\{ \lambda^2 + \lambda - \sum_{\alpha\in J} E X_\alpha W \right\}$$

$$= (1\wedge\lambda^{-1}) \left\{ \lambda^2 + \lambda - E W^2 \right\}$$

$$= (1\wedge\lambda^{-1}) \left\{ \lambda - \mathrm{Var}(W) \right\}$$

$$= (1\wedge\lambda) \left\{ 1 - \frac{\mathrm{Var}(W)}{\lambda} \right\}$$

this proves (4.7)

Under PR,

$$\left|\sum_{\alpha\in J} p_\alpha E\left[f_A(W+1) - f_A(V_\alpha+1)\right]\right|$$

$$\leq \left|\sum_{\alpha\in J} p_\alpha \left[f_A(W+1) - f_A(W^{(\alpha)}+1)\right]\right|$$

$$+ \left|\sum_{\alpha\in J} p_\alpha \left[f_A(W^{(\alpha)}+1) - f_A(V_\alpha+1)\right]\right.$$

$$\leq \sum_{\alpha \in J} (n\lambda^{-1}) p_\alpha E X_\alpha$$

$$+ \sum_{\alpha \in J} (n\lambda^{-1}) p_\alpha E | V_\alpha - W^{\alpha} |$$

$$= (n\lambda^{-1}) \sum_{\alpha \in J} p_\alpha^2$$

$$+ (n\lambda^{-1}) \sum_{\alpha \in J} p_\alpha E ((V_\alpha + 1) - (W^{\alpha} + 1))$$

(since $X_{\beta\alpha} \geq X_\beta$)

$$= (n\lambda^{-1}) \sum_{\alpha \in J} p_\alpha^2$$

$$+ (n\lambda^{-1}) \sum_{\alpha \in J} p_\alpha E(W^{\alpha} + 1 | X_\alpha = 1)$$

$$- (n\lambda^{-1}) \sum_{\alpha \in J} \{ p_\alpha (\lambda - p_\alpha) - p_\alpha \}$$

$$= (n\lambda^{-1}) \sum_{\alpha \in J} p_\alpha^2$$

$$+ (n\lambda^{-1}) \sum_{\alpha \in J} E X_\alpha W$$

$$- (n\lambda^{-1}) \{ \lambda^2 - \sum_{\alpha \in J} p_\alpha^2 - \lambda \}$$

$$= (n\lambda) \left\{ \frac{Var(W)}{\lambda} - 1 + \frac{2}{\lambda} \sum_{\alpha \in J} p_\alpha^2 \right\}.$$

This proves (4.8)

Application

(1) Classical occupancy problem

Throw r balls independently into n boxes with prob. $\theta_1, \theta_2, \dots, \theta_n$ $\left(\sum_{\alpha=1}^{n} \theta_\alpha = 1\right)$

What is the distribution of the number of empty boxes?

Take $J = \{1, \dots, n\}$.

Define $X_\alpha = I(\alpha^{th}$ box is empty$)$.

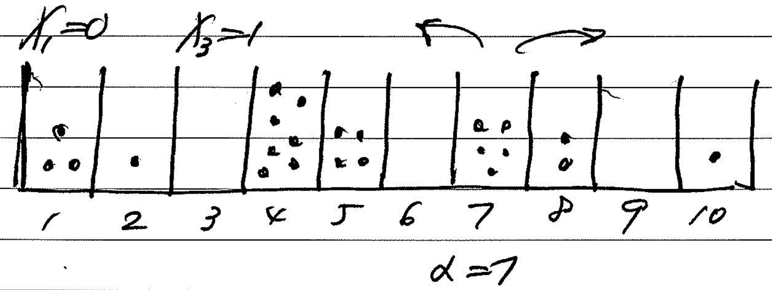
Then $W = \sum_{\alpha \in J} X_\alpha$ = number of empty boxes

$$P(X_\alpha = 1) = 1 - P(X_\alpha = 0) = (1-\theta_\alpha)^r,$$

$$\lambda = EW = \sum_{\alpha=1}^{n} (1-\theta_\alpha)^r.$$

Define $Z \sim Po(\lambda)$.

For each $\alpha \in J$, construct $\{Y_{\beta\alpha} : \beta \in J\}$ as follows:



If the αth box is empty, define $Y_{\beta\alpha} = X_\beta$ $\forall \beta \in J$.

Otherwise, throw all the balls at the α^{th} box

independently into the other boxes with prob. $\frac{\theta_\beta}{1-\theta_\alpha}$
for $\beta \neq \alpha$. Define $Y_{\beta\alpha} = I(\beta^{th}$ box is empty$)$.

Then $\mathcal{L}(Y_{\beta\alpha} : \beta \in J) = \mathcal{L}(X_\beta : \beta \in J \mid X_\alpha = 1)$

$\quad$ & $Y_{\beta\alpha} \le X_\beta \quad \forall \beta \in J. \quad \Rightarrow NR$

So $d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda)\left\{1 - \frac{Var(W)}{\lambda}\right\}$

$$Var(W) = EW^2 - \lambda^2 = \sum_{\alpha \neq \beta}\sum E X_\alpha X_\beta + \lambda - \lambda^2$$

$$= \sum_{\alpha \neq \beta}\sum (1-\theta_\alpha-\theta_\beta)^r + \lambda - \lambda^2$$

So $d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda)\left\{\lambda - \frac{1}{\lambda}\sum_{\alpha \neq \beta}\sum (1-\theta_\alpha-\theta_\beta)^r\right\}$

(2) <u>Sampling without replacement</u>

N objects of which m are of one category & $N-m$ of another category.

Sample n objects without replacement.

Let W = number of objects of 1^{st} category in the sample.

W has a hypergeometric distribution.

No

Date

48

Same experiment as the following:

Arrange the N object at random.

W = number of objects of the 1st category at positions $1, \dots, n$.

$$P(W=j) = \frac{\binom{m}{j}\binom{N-m}{n-j}}{\binom{N}{n}} = \frac{\binom{n}{j}\binom{N-n}{m-j}}{\binom{N}{m}}.$$

Take $J = \{1, \dots, n\}$

$$\lambda = EW = \frac{nm}{N}$$

$$\operatorname{Var}(W) = \frac{N-n}{N-1}\frac{nm}{N}\left(1-\frac{m}{N}\right)$$

Define $X_\alpha = I$(object of 1st category at position α)

Then $W = \sum_{\alpha \in J} X_\alpha$. Define $Z \sim Po(\lambda)$.

For each $\alpha \in J$, construct $\{Y_{\beta\alpha} : \alpha \in J\}$ as follows:

If $X_\alpha = 1$, define $Y_{\beta\alpha} = X_\beta$.

If $X_\alpha = 0$, select an object of 1st category randomly & switch it with the object of 2nd category at α. Then define $Y_{\beta\alpha} = I$(object of 1st category at β)

Then $\mathcal{L}(Y_{\beta\alpha} : \beta \in J) = \mathcal{L}(X_\beta : \beta \in J \mid X_\alpha = 1)$

& $Y_{\beta\alpha} \le X_\beta$ for $\beta \in J$. $\Rightarrow$ NR

So $d(\mathcal{L}(W), \mathcal{L}(Z)) \le (1 \wedge \lambda)\left\{1 - \frac{\operatorname{Var}(W)}{\lambda}\right\}$

$$= (\Gamma\wedge\lambda)\left\{1 - \frac{\frac{N-n}{N-1}\frac{nm}{N}\left(1-\frac{m}{N}\right)}{\frac{nm}{N}}\right\}$$

$$= (\Gamma\wedge\lambda)\left\{1 - \frac{N-n}{N-1}\left(1-\frac{m}{N}\right)\right\}$$

$$= (\Gamma\wedge\lambda)\frac{N}{N-1}\left(\frac{n}{N}+\frac{m}{N}-\frac{nm}{N^2}-\frac{1}{N}\right).$$

Now Vatutin & Mikhailov (Theory Probab. Appl 1982):

W has the same distribution as
a sum of independent indicators.
By Theorem 3.2 & (4.5)

$$d(\mathcal{L}(W), \mathcal{L}(Z)) \le (\Gamma\wedge\lambda)\left\{1 - \frac{\mathrm{Var}(W)}{\lambda}\right\}$$

$$= (\Gamma\wedge\lambda)\frac{N}{N-1}\left(\frac{n}{N}+\frac{m}{N}-\frac{nm}{N^2}-\frac{1}{N}\right)$$

Same bound.

(3) <u>A random graph problem</u>

K_n complete graph with n vertices & $\binom{n}{2}$ edges,
G a given subgraph with $v(G)$ vertices
and $e(G)$ edges.

Assume $e(G) >$ and G has no isolated points.

Delete each edge of K_n with prob $1-\theta$ independently of other edges. Thus get a random graph $K_{n,\theta}$

How many subgraphs of $K_{n,\theta}$ are isomorphic to G?

Take J = the set of all copies of G in K_n.

$\forall \alpha \in J$, define $X_\alpha = I(\alpha \in K_{n,\theta})$

& $W = \sum_{\alpha \in J} X_\alpha$.

Then W = number of subgraphs of $K_{n,\theta}$ which are isomorphic to G.

$$p_\alpha = P(X_\alpha = 1) = \theta^{e(G)}$$

(The assumption of no isolated points is relevant here).

$$\lambda = EW = \binom{n}{v(G)} \frac{(v(G))!}{a(G)} \theta^{e(G)}$$

where $a(G) = |$ Automorphism group of $G|$

= number of permutations of the set of vertices of G which leave G invariant.

Let $Z \sim Po(\lambda)$.

Fix α. $\forall \beta \in J$, define $Y_{\beta\alpha} = I(\beta \in K_{n,\theta} \cup \alpha)$

Clearly $Y_{\beta\alpha} \geq X_\beta$ $\forall \beta \in J$. $\Rightarrow$ PR

By independence in $K_{n,\theta}$,

$$\mathcal{L}(Y_{\beta\alpha} : \beta \in J) = \mathcal{L}(X_\beta : \beta \in J \mid X_\alpha = 1)$$

So $$d(\mathcal{L}(W), \mathcal{L}(Z)) \leq (1 \wedge \lambda)\left\{ \frac{\mathrm{Var}(W)}{\lambda} - 1 + 2\theta^{e(G)} \right\}$$

$$\leq \sum_H \frac{a(H)}{a(G)} c(G,H)\, n^{v(G)-v(H)} \theta^{e(G)-e(H)}$$

where $c(G,H)$ is the number of copies of H in G and the sum is over all non-empty subgraphs $H \subsetneq G$ without isolated vertices.

Find conditions for which the bound is small ($d \to 0$ as $n \to \infty$).

§5 Existence of monotone couplings

$S = \{0,1\}^n$

$x = (x_1, \dots, x_n) \in S$

$y = (y_1, \dots, y_n) \in S$

Definitions

(i) $y \le x$ if $y_i \le x_i$ for all $i = 1, \dots, n$.

(ii) $\varphi : S \to \mathbb{R}$ is increasing if

$\varphi(y) \le \varphi(x)$ for all $y \le x$.

(iii) $\varphi : S \to \mathbb{R}$ is decreasing if

$-\varphi$ is increasing

(iv) Consider random element (I, J) on $S \times S$

having distribution μ with given

marginals $\mathcal{L}(I)$ & $\mathcal{L}(J)$. If such

a probability measure μ exists

and fulfils $\mu(I \ge J) = 1$,

then we say, there exists a monotone

coupling of I & J with $I \ge J$.

Theorem 5.1

A coupling with $I \geq J$ exists if and only if for every increasing indicator function φ

$$E\varphi(I) \geq E\varphi(J).$$

(A proof can be found on p. 72 of Liggett, T. M. Interacting Particle Systems, Springer, New York, 1985)

Corollary 5.2

Suppose that ψ is an indicator function s.t. $P(\psi(I)=1)>0$ and $\mathcal{L}(J) = \mathcal{L}(I \mid \psi(I)=1)$. Then a coupling with $I \geq J$ ($I \leq J$) exists if and only if $\psi(I)$ and $\varphi(I)$ are negatively (positively) correlated for every increasing indicator function φ.

Proof We prove negative correlation. From defn of J,

$$E\varphi(J) = E(\varphi(I) \mid \psi(I)=1)$$

$$= \frac{E\,\psi(I)\,\varphi(I)}{P(\psi(I)=1)}$$

$$= \frac{E\psi(I)\varphi(I)}{E\psi(I)}$$

By theorem 5-1,

$$I \geq J \iff E\varphi(I) \geq E\varphi(J)$$

$$\iff E\varphi(I) \geq \frac{E\psi(I)\varphi(I)}{E\psi(I)}$$

$$\iff E\varphi(I)E\psi(I) \geq E\psi(I)\varphi(I).$$

Proposition 5-3

Any set $X_1, \dots, X_n$ of independent random variables satisfy the FKG inequality: if f & g are bounded increasing functions, then

$$Ef(X)g(X) \geq Ef(X)Eg(X).$$

where $X = (X_1, \dots, X_n)$.

(A proof is given on p.76 of Liggett (1985)).

Exercise: Prove that for one r.v. X, the FKG inequality always holds.

Theorem 5.4

$\{ I_\alpha : \alpha \in \Gamma \}$ increasing indicator functions of independent random variables $X_1, \dots, X_n$. Then $\{ I_\alpha : \alpha \in \Gamma \}$ are positively related.

Proof.

Take $I = (I_\beta : \beta \in \Gamma)$

and $\forall \alpha \in \Gamma$, let $\psi(I) = I_\alpha$.

Then the FKG inequality implies that $\forall$ bounded increasing function φ

$$E I_\alpha \varphi(I) \geq E I_\alpha E \varphi(I)$$

Since I_α is also a bdd increasing function of I.

By corollary 5.2, for any J s.t.

$$\mathcal{L}(J) = \mathcal{L}(I_\beta : \beta \in \Gamma \mid I_\alpha = 1)$$

we have the coupling $J \geq I$.

That is $\{ I_\alpha : \alpha \in \Gamma \}$ are positively related.

Definition

Any collection $X_1, \ldots, X_n$ of random variables are said to be associated if they satisfy the FKG inequality.

(By Prop. 5.3, indep r.v.'s are associated)

Proposition 5.5

Random variables which are increasing functions of associated random variables are associated.

Proof It follows immediately from definition

Theorem 5.6

Let $\{X_\nu\}$ be associated random variables. Suppose $\forall \alpha \in \Gamma$, I_α is an increasing (decreasing) indicator function of every X_ν.
Then $\{I_\alpha : \alpha \in \Gamma\}$ are positively related.

Proof same as for Theorem 5.4.

No

Date

(4) Exceedance of stationary sequences

Suppose $\{\xi_i\}_{-\infty}^{\infty}$ are i.i.d. & $\{c_i\}_{-\infty}^{\infty}$ nonnegative numbers s.t. $\eta_k = \sum_{i=-\infty}^{\infty} c_i \xi_{k-i}$ converges a.s. Then $\{\eta_k\}_{-\infty}^{\infty}$ is a stationary sequence.

Interested in

(i) Number of η_k which exceed a certain value;

(ii) Asymptotic distribution of $\max_{1\le k\le n} \eta_k$ as $n \to \infty$.

Let $X_k = I(\eta_k > z)$.

then $W = \sum_{k=1}^{n} X_k$ is the number of exceedances above the level z among $\eta_1, \dots, \eta_n$.

$\{W=0\} = \{\max_{1\le k\le n} \eta_k \le z\}$.

Now each X_k is an increasing indicator function of the associated r.v.'s $\{\xi_i\}_{-\infty}^{\infty}$.

$\therefore$ $X_1, X_2, \dots X_n$ are positively related.

$P(X_k=1) = 1 - P(X_k=0) = P(\eta_k > z) = P(\eta_0 > z)$.

$\lambda = EW = nP(\eta_0 > z)$. Let $Z \sim Po(\lambda)$.

then $|P(\max_{1\le k\le n} \eta_k \le z) - e^{-\lambda}| \le d(\mathcal{L}(W), \mathcal{L}(Z))$

$\le (1 \wedge \lambda)\left\{\frac{Var(W)}{\lambda} - 1 + 2P(\eta_0 > z)\right\}$

Definition

A collection $\{X_\alpha\}$ of random variables is said to be negatively associated if for all disjoint subsets A_1 & A_2 of indices and all bounded functions f & g increasing in every variable,

$$E f(X_\alpha : \alpha \in A_1)\, g(X_\alpha : \alpha \in A_2)$$

$$\leq E f(X_\alpha : \alpha \in A_1)\, E g(X_\alpha : \alpha \in A_2)$$

Theorem 5.7

Let $\{X_\nu\}$ be negatively associated r.v.s and $\{S_\alpha\}$ disjoint subsets of them. Suppose $\forall \alpha \in \Gamma$ that $I_\alpha = I_\alpha(S_\alpha)$ is an increasing (decreasing) indicator function of every $X_\nu \in S_\alpha$. Then the random variables $\{I_\alpha : \alpha \in \Gamma\}$ are negatively related.

In particular, negatively associated indicators are negatively related.