Theory 1 Practice Exam 2012

1). $N \sim Poi(\mu)$

$X_1, \dots, X_n \overset{iid}{\sim} Poi(\lambda)$

$0 < \mu, \lambda < \infty$

$U = I(N > 0) \sum_{i=1}^{N} X_i$

(a) Show $E[U] = \mu\lambda$ and $Var[U] = \mu\lambda(1+\lambda)$

$U | N \sim Poi(N\lambda)$

$E[U] = E[E[U|N]] = E[N\lambda]$

$= \lambda E[N] = \mu\lambda$ ✓

$Var[U] = Var[E[U|N]] + E[Var[U|N]]$

$= Var[N\lambda] + E[N\lambda]$

$= \lambda^2 Var[N] + \lambda E[N]$

$= \lambda^2 \mu + \lambda\mu$

$= \lambda\mu(1+\lambda)$ ✓

* Could also write out $E[E[I(N>0)\sum_{i=1}^{N} X_i | N]] \dots$

$E[E[I(N>0)\sum_{i=1}^{N} X_i | N]] = E[N E(X_i)] = E[N\lambda] = \lambda\mu$

(b) Now:

$N \sim Poi(K)$

$X_1, \ldots, X_N \sim Poi(\lambda_K = h/K)$ $\quad 0 < h < \infty$, fixed

What happens to U as $K \to \infty$?

$D_i = I(X_i = 1)$ for $i \geq 1$

$$T \stackrel{\cdot}{=} I(N > 0) \sum_{i=1}^{N} D_i$$

(i) Derive the limits of $E[U]$ & $Var[U]$ as $K \to \infty$

$$E[U] = \mu\lambda = K(h/K) = h$$

$$\lim_{K \to \infty} E[U] = h$$

$$Var[U] = \mu\lambda(1+\lambda)$$

$$= K\left(\frac{h}{K}\right)\left(1 + \frac{h}{K}\right)$$

$$= h\left(1 + \frac{h}{K}\right)$$

$$\lim_{K \to \infty} Var[U] = \lim_{K \to \infty} h\left(1 + \frac{h}{K}\right) = h$$

typo in test

$\frac{1}{2}$ $o(1)$

(ii) Show that $P(x_i \neq D_i) = \lambda_k^2 (1 + o(\lambda_k))$ as $k \to \infty$

$$P(x_i \neq D_i) = 1 - P(x_i = D_i)$$
$$= 1 - P(D_i = x_i = 1 \text{ or } D_i = x_i = 0)$$
$$= 1 - [P(x_i = 1) + P(x_i = 0)]$$
$$= 1 - [\lambda_k e^{-\lambda_k} + e^{-\lambda_k}]$$
$$= 1 - e^{-\lambda_k}[\lambda_k + 1]$$

$$e^{\mu} = \sum_{n=0}^{\infty} \frac{\mu^n}{n!}$$

$$= 1 - \left[1 + (-\lambda_k) + \frac{\lambda_k^2}{2} - \frac{\lambda_k^3}{6} + \dots\right][\lambda_k + 1]$$

$$\cancel{1} - [\cancel{1} - \lambda_k + \lambda_k^2/2 - \lambda_k^3/6 + \dots] +$$
$$- \lambda_k[1 - \lambda_k + \lambda_k^2/2 - \lambda_k^3/6 + \dots]$$
$$= \cancel{\lambda_k} - \frac{\lambda_k^2}{2} + \frac{\lambda_k^3}{6} + [R] - \cancel{\lambda_k} + \lambda_k^2 - \frac{\lambda_k^3}{2} + \frac{\lambda_k^4}{6} + \lambda_k[R]$$

$$= \lambda_k^2[1 - 1/2 + \lambda_k/6 - \lambda_k/2 + \lambda_k^2/24 + \lambda_k^2/6 + \dots]$$
$$= \lambda_k^2\left[1/2 + \sum_{t=1}^{\infty} \lambda_k^t \left(\frac{1}{(t+2)!} - \frac{1}{(t+1)!}\right)\right]$$
$$\hookrightarrow (1-(t+2))/(t+2)! = \frac{1-t}{(t+2)!}$$

Defn: $o(\lambda_k) \Rightarrow \frac{1}{\lambda_k}|x_k| \to 0$

$$\cancel{= \lambda_k^2\left[1/2 + \lambda_k \sum_{t=1}^{\infty} \lambda_k^{t-1}\left(\frac{1}{(t+2)!} - \frac{1}{(t+1)!}\right)\right]}$$

$$\lambda_k o(1) = o(\lambda_k)$$

$$\lambda_k = h/k \Rightarrow \lim_{k\to\infty} \lambda_k^t = \lim_{k\to\infty}\left(\frac{h}{k}\right)^t = 0$$

$$\Rightarrow \sum_{t=1}^{\infty} \lambda_k^t \left(\frac{1}{(t+2)!} - \frac{1}{(t+1)!}\right) = o(1) \checkmark$$

(iii) Show:

$$I(U \neq T) \leq I(N > 0) \sum_{i=1}^{N} I(x_i \neq D_i)$$

Need to show whenever right side $= 0 \Rightarrow$ the left side also $= 0$.

- If right side $> 0 \Rightarrow$ right side 1 or more $\Rightarrow$ expression always true.

Right side $= 0$:

$*\ N = 0$ OR $*\ x_i = D_i \quad \forall i = 1, \dots, N$

$\hookrightarrow x_i = D_i = 0, \quad x_i = D_i = 1$

<u>Case 1:</u> $N = 0$, no restriction on $x_i + D_i$

$U = I(N > 0) \sum_{i=1}^{N} x_i$

$= 0$

$T = I(N > 0) \sum_{i=1}^{N} D_i$

$= 0$

$\Rightarrow I(U \neq T) = 0$ since $U = T$ ✓

<u>Case 2:</u> $N > 0$, $x_i = D_i = 0$ OR $x_i = D_i = 1 \quad \forall i = 1, \dots, N$

$\Rightarrow \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} D_i$

$\Rightarrow I(N > 0) \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} D_i \, I(N > 0)$

$\Rightarrow U = T$

$\Rightarrow I(U \neq T) = 0$ ✓

Show $U - T \xrightarrow{P} 0$ as $K \to \infty$

If statement is true, then

$E[I(U \neq T)] \leq E[I(N > 0) \sum_{i=1}^{N} I(x_i \neq 0_i)]$

$\Rightarrow P(U \neq T) \leq$ (above)

$\Rightarrow P(|U - T| > \varepsilon) \leq$ (above) for $\varepsilon > 0$

If we can show $E[I(N > 0) \sum_{i=1}^{N} I(x_i \neq 0_i)] \to 0$ as $K \to \infty$,

then $P(|U - T| > \varepsilon) \to 0$

$\Rightarrow U - T \xrightarrow{P} 0$

$E[I(N > 0) \sum_{i=1}^{N} I(x_i \neq 0_i)]$

$= E[E[I(N > 0) \sum_{i=1}^{N} I(x_i \neq 0_i) \mid N]]$

$= E[N E[I(x_i \neq 0_i)]]$

expression $| N \sim \text{Bin}(N, P(x_i \neq 0_i))$

$= E[N P(x_i \neq 0_i)]$

$= \mu_K P(x_i \neq 0_i)$

$= \mu_K \lambda_K^2 (1 + o(\lambda_K))$

$= K \left(\frac{h}{K}\right)^2 (1 + o(\lambda_K))$

$= h\left(\frac{h}{K}\right)\left(1 + o\left(\frac{h}{K}\right)\right)$

If $x_K = o(h/K) \Rightarrow \frac{|x_K|}{(h/K)} \to 0 \Rightarrow \frac{K|x_K|}{h} \to 0$

$\Rightarrow \left(\frac{h}{K}\right)\left(\frac{K}{h}|x_K|\right)$ also $\to 0$

$\lim_{K \to \infty} \frac{h^2}{K}(1 + o(h/K)) \to 0$ ✓

(iv) Show that $T - \sum_{i=1}^{K} D_i \xrightarrow{P} 0$ as $K \to \infty$

$$T = I(N>0) \sum_{i=1}^{N} D_i$$

$$T - d = I(N>0) \sum_{i=1}^{N} D_i - \sum_{i=1}^{K} D_i = \cancel{\sum_{i=1}^{K} D_i \left(I(N>0) - 1\right)}$$

$$P(|T-d| > \varepsilon) = P\left(\left|I(N>0) \sum_{i=1}^{N} D_i - \sum_{i=1}^{K} D_i\right| > \varepsilon\right) =$$

Want to show above $\to 0$ OR find an ~~sequence~~ expression it is $\leq$ and show this expression $\to 0$

$N \sim Poi(K)$

Show $I(N>0) - 1 \xrightarrow{P} 0$ as $K \to \infty$

$\Rightarrow I(N>0) \xrightarrow{P} 1$ as $K \to \infty$

$\Rightarrow I(N=0) \xrightarrow{P} 0$

$\Rightarrow P(N=0) \xrightarrow{K\to\infty} 0$

$P(N=0) = e^{-\mu_K}$

$= e^{-K}$

$\lim_{K\to\infty} e^{-K} = 0$ ✓

Hint: (iii) not helpful here

Study the difference.

Cases: $N > K$, $K > N$

(v) uses (iii) + (iv)

$\sum_{i=1}^{K} D_i \sim Bin(K, \_)$ for (v)

(iv) Show that $T - \sum_{i=1}^{K} D_i \xrightarrow{P} 0$ as $K \to \infty$

$$T - \sum_{i=1}^{K} D_i = I(N>0) \sum_{i=1}^{N} D_i - \sum_{i=1}^{K} D_i$$

Case 1: $N > K$

$$\left| T - \sum_{i=1}^{K} D_i \right| = \sum_{i=K+1}^{N} D_i$$

Case 2: $N < K$

$$\left| T - \sum_{i=1}^{K} D_i \right| = \sum_{i=N+1}^{K} D_i$$

Case 3: $N = K$

$$\left| T - \sum_{i=1}^{K} D_i \right| = 0$$

see alternative method in a few pages as well

Want to show $P(|T - \sum_{i=1}^{K} D_i| > \varepsilon) \to 0$ as $K \to \infty$

$$P(|T - \sum_{i=1}^{K} D_i| > \varepsilon) = P(|T - \sum_{i=1}^{K} D_i| \neq 0)$$

$$= P(|\sum_{i=K+1}^{N} D_i| > 0) P(N>K) + P(|\sum_{i=N+1}^{K} D_i| > 0) P(N<K)$$

$$= (1 - P(|\sum_{i=K+1}^{N} D_i| = 0)) P(N>K) + (1 - P(|\sum_{i=N+1}^{K} D_i| = 0)) P(N<K)$$

$$= (1 - P(x_{K+1} \neq 1, \ldots, x_N \neq 1)) P(N>K) + (1 - P(x_{N+1} \neq 1, \ldots, x_K \neq 1)) P(N<K)$$

$$= (1 - (P(x_i \neq 1))^{N-K}) P(N>K) + (1 - (P(x_i \neq 1))^{K-N}) P(N<K)$$

$$= (1 - (1 - P(x_i = 1))^{N-K}) P(N>K) + (1 - (1 - P(x_i = 1))^{K-N}) P(N<K)$$

$$= (1 - (1 - \lambda_K e^{-\lambda_K})^{N-K}) P(N>K) + (1 - (1 - \lambda_K e^{-\lambda_K})^{K-N}) P(N<K)$$

$$= \underbrace{\left(1 - \frac{(1 - (h/K) e^{-h/K})^N}{\left(1 - \frac{h e^{-h/K}}{K}\right)^K} \right)}_{(*)} P(N>K) + \underbrace{\left(1 - \frac{\left(1 - \frac{h e^{-h/K}}{K}\right)^K}{(1 - (h/K) e^{-h/K})^N} \right)}_{(**)} \cdot P(N<K)$$

(*)
$$\lim_{K \to \infty} \frac{(1 - (h/K) e^{-h/K})^N}{\left(1 - \frac{h e^{-h/K}}{K}\right)^K}$$

$$= \lim_{K \to \infty} (1 - (h/K) e^{-h/K})^N \lim_{K \to \infty} \frac{1}{\left(1 - \frac{h e^{-h/K}}{K}\right)^K}$$

$\longrightarrow$

Note:

$$\lim_{K\to\infty} (1 + A_K)^K = e^A \quad \text{for } A_K \xrightarrow{K\to\infty} A$$

Also: since $N > K \Rightarrow N \to \infty$

$$= e^{-h} \cdot \frac{1}{e^{-h}} = 1$$

** $\lim_{K\to\infty} \dfrac{\left(1 - \frac{he^{-h/K}}{K}\right)^K}{(1-(h/K)e^{-h/K})^N} = e^{-h}(1)$

$N < K$: $1-(h/K)e^{-h/K} \to 1 - 0 = 1$

*** $P(N > K) = 1 - P(N \le K)$

$$= 1 - \sum_{i=0}^{K} \frac{M_K^{\,i} e^{-K}}{i!}$$

$$= 1 - \sum_{i=0}^{K} \frac{K^i e^{-K}}{i!}$$

$\lim_{K\to\infty}$ (above) $= 1 - 1 = 0$ (?)

$$P(K > N) = \sum_{N=0}^{K-1} \frac{(K)^i e^{-K}}{i!}$$

$$= \sum_{N=0}^{K-1} \frac{(K)^i e^{-K}}{i!}$$

N considered fixed, $K \to \infty$

$$\lim_{K\to\infty} P(K > N) = 1$$

Now we have

$(1 - e^{-h}/e^{-h})(0) + (1 - e^{-h})(1)$

$= 0 + 0$

$= 0$ ✓

Note: Probabilities need to be slightly changed:

$P(|\Sigma_{i=K+1}^{N} D_i| > 0 \mid N > K)$ and

$P(|\Sigma_{i=N+1}^{K} D_i| > 0 \mid \underbrace{N < K}_{\text{need given part to be clear}})$

Probabilites need to be slightly fixed:

$P(|\Sigma_{i=k+1}^{N} D_i| > 0 \mid N > k)$ and

$P(|\Sigma_{i=N+1}^{k} D_i| > 0 \mid N < k)$

Alternative method:

$$E[\,|T - \Sigma_{i=1}^{k} D_i| \mid N] = |N-k|\, E[D_i \mid N]$$

$$\Rightarrow E[E[\,|T - \Sigma_{i=1}^{k} D_i| \mid N]] = E[\,|N-k|\, E[D_i \mid N]]$$

$$= E[\,|N-k|]\, E[D_i]$$

explanation below

$$\leq \sqrt{k}\, \lambda_k e^{-\lambda_k} \qquad \lambda_k = h/k$$

$$E[(N-k)^2] = Var(N) = k$$

$$= \sqrt{k} \left(\frac{h}{k}\right) e^{-h/k}$$

$$= \frac{h}{k^{1/2}} e^{-h/k}$$

$$\lim_{k \to \infty} (\text{above}) = 0$$

$$P(|T - \Sigma_{i=1}^{k} D_i| > \varepsilon) \leq \frac{E[\,|T - \Sigma_{i=1}^{k} D_i|\,]}{\varepsilon} \longrightarrow 0$$

Markov's Inequality

$$E[\,|N-k|\,] \leq E[(N-k)^2]^{1/2}$$ by Holder's Inequality

$$= (k)^{1/2}$$

since $E[(N-k)^2] = Var(N) = k$

(v) Show U converges in distribution to a Poisson RV w/ parameter h as $k \to \infty$

$$\sum_{i=1}^{k} D_i \sim \text{Bin}(k, P(X_i = 1))$$
$$\sim \text{Bin}\left(k, \frac{h}{k} e^{-h/k}\right)$$

We know that for large n and small p, $\text{Bin}(n,p)$ approximates the Poisson distribution.

$$\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{\lambda^x e^{-\lambda}}{x!} \text{ for } \lambda = np.$$

Consequently,

$$\lim_{k\to\infty} \binom{k}{y} \left(\frac{h}{k} e^{-h/k}\right)^y \left(1 - \frac{h}{k} e^{-h/k}\right)^{k-y} =$$

$$\frac{\lim_{k\to\infty} \left(k\left(\frac{h}{k} e^{-h/k}\right)\right)^y e^{-k(h/k)e^{-h/k}}}{y!}$$

$$= \frac{(h(1))^y e^{-h(1)}}{y!}$$

$$= \frac{h^y e^{-h}}{y!}$$

$\sim$ Poisson(h) ✓

by (iv), we know $T - \sum_{i=1}^{k} D_i \xrightarrow{P} 0$

$\Rightarrow$ ~~$T \xrightarrow{P} \sum_{i=1}^{k} D_i$~~ (dist of)

by (iii), we know $U - T \xrightarrow{P} 0$

$\Rightarrow$ ~~$U \xrightarrow{P} T$~~

next pg

$\Rightarrow U \xrightarrow{P} (\sum_{i=1}^{k} D_i)\ \text{Poi}(h)$

$\Rightarrow U \xrightarrow{d} \text{Poi}(h)$ ✓

by (ii), $T - \sum_{i=1}^{k} D_i \xrightarrow{p} 0$

and by (iii) $U - T \xrightarrow{p} 0$

$\Rightarrow (U - T) + (T - \sum_{i=1}^{k} D_i) \xrightarrow{p} 0$

$\Rightarrow U - \sum_{i=1}^{k} D_i \xrightarrow{p} 0$

$\Rightarrow U - \sum_{i=1}^{k} D_i \xrightarrow{d} 0$

$\Rightarrow U \xrightarrow{d}$ asymptotic dist of $\sum_{i=1}^{k} D_i$

$\Rightarrow U \xrightarrow{d} Poi(h)$

Alternatively, could say

$U = \sum_{i=1}^{k} D_i + o_p(1)$

Therefore, by Slutsky's thm,

$U \equiv \sum_{i=1}^{k} D_i + o_p(1) \xrightarrow{d} \text{Asym. Dist}(\sum_{i=1}^{k} D_i)$ ✓

(v) Show $U \xrightarrow{d} Poi(h)$ as $K \to \infty$

- Could find MGF / characteristic function & show it converges to that of a $Poi(h)$ RV
- Could show $F_K(u) \xrightarrow{d} \underbrace{F(u)}_{Poi(h)}$ X

$$M_U(t) = E[e^{tU}]$$
$$= E[E[e^{tU} \mid N]]$$

given N, $U \sim Poi(N\lambda_K)$

$$= E[\exp(N\lambda_K(e^t - 1))]$$

$N \sim Poi(\mu_K)$

$$= E\left[\sum_{j=0}^{\infty} \frac{(N\lambda_K(e^t-1))^j}{j!}\right]$$
$$= \sum_{j=0}^{\infty} \frac{(\lambda_K(e^t-1))^j}{j!} E[N^j]$$

$E[N] = \mu_K = K$

$E[N^2] = Var(N) + (E(N))^2 = \mu_K + \mu_K^2 = K(K+1) = K^2 + K$

$E[N^3] = \mu_K(\mu_K+1)^2 + \mu_K^2 = \mu_K^3 + 3\mu_K^2 + \mu_K =$
$= K^3 + 3K^2 + K$

By some process (mathematical induction...?), we can show $E[N^j]$ = polynomial of degree j

$$\frac{E[N^j]}{K^j} = 1 + o_p(1)$$

$$= \sum_{j=0}^{\infty} \frac{(h/K)^j (e^t-1)^j E[N^j]}{j!}$$
$$= \sum_{j=0}^{\infty} \frac{E[N^j]}{K^j} \frac{(h(e^t-1))^j}{j!}$$
$$= \sum_{j=0}^{\infty} (1 + o_p(1)) \frac{(h(e^t-1))^j}{j!}$$

Continued $\longrightarrow$

$$\lim_{K\to\infty} \sum_{j=0}^{\infty} (1+op(1))\frac{(\mu(e^t-1))^j}{j!}$$

$$= \sum_{j=0}^{\infty} \frac{(\mu(e^t-1))^j}{j!}$$

$$= \exp(\mu(e^t-1))$$

$= \text{MGF of } Poi(\mu)$ ✓

Cumulant generating function:

$$K_N(t) = \log M_N(t) = \mu(e^t-1)$$

$$\frac{\partial^j}{\partial t^j}\mu(e^t-1)\Big|_{t=0} = \mu(e^t)\Big|_{t=0} = \mu \quad \forall j$$

$= \text{central moment}$

$$= E[(N-\mu)^j]$$

$$= E\left[\sum_{l=0}^{j} N^{j-l}\mu^j\binom{j}{l}\right]$$

$$= \binom{j}{l}\sum_{l=0}^{j}\mu^j E[N^{j-l}] = \mu$$

$$= E[N(N-\mu)^j - \mu(N-\mu)^j]$$

$$= E[N(N-\mu)^j] - \mu(\mu) = \mu$$

$$\Rightarrow E[N(N-\mu)^j] = \mu + \mu^2$$

... ?

(c) $N \sim Poi(h/k)$ $\mu_k = h/k$

$X_i \sim Poi(k)$ $\lambda_k = k$

(i) $E[U] = \lambda_k \mu_k$

$= k(h/k)$

$= h$

$\lim_{k \to \infty} E[U] = h$

$Var[U] = \lambda_k \mu_k (1 + \lambda_k)$

$= k\left(\frac{h}{k}\right)(1+k)$

$= h(1+k)$

$\lim_{k \to \infty} Var[U] = \infty \rightarrow$ limit DNE

(ii) Show $U \xrightarrow{d} 0$ as $k \to \infty$ (bound by N)

Cases: $N > 0$, $N = 0$

Show $U \xrightarrow{P} 0 \Rightarrow U \xrightarrow{d} 0$

$P(|U - 0| > \varepsilon) \le \frac{E[(U-0)^2]}{\varepsilon^2}$ $\longrightarrow$

$= \frac{Var(U) + (E[U])^2}{\varepsilon^2}$

$= \frac{h(1+k) + h^2}{\varepsilon^2}$

(ii) Show that $U \to 0$ in dist as $K \to \infty$

Show $F(u) \to I(u \geq 0)$

- at 0, jumps from 0 to 1

OR:

$P(U=0) = E[P(U=0 \mid N)]$

$= E[P(X_1=0)^N]$

$= \ldots$

replace $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$

$= e^{-\mu k} e^{\mu k e^{-\lambda k}}$

$\to 1$

$\Rightarrow P(U=0) \to 1$ as $k \to \infty$ $\Rightarrow U \to 0$ as $k \to \infty$

$P(U=0) = P(U=0 \mid N=0) P(N=0) + P(U=0 \mid N>0) P(N>0)$

$= \ldots$

OR:

$P(U \leq u) = P(I(N>0) \sum_{i=1}^{N} X_i \leq u)$

$= P(\sum_{i=1}^{N} X_i \leq u \mid N>0) P(N>0) +$

$+ P(\sum_{i=1}^{N} X_i \leq u \mid N=0) P(N=0)$

$= P(\sum_{i=1}^{N} X_i \leq u \mid N>0) P(N>0) +$

$+ P(0 \leq u \mid N=0) P(N=0)$

$= O(1) \overbrace{(1-e^{-\mu k})}^{o(1)} + I(u \geq 0) e^{-\mu k}$

$= o(1) + I(u \geq 0)(1+o(1))$

$\to I(u>0)$ as $k \to \infty$

(ii) Show $U \xrightarrow{d} 0$ as $k \to \infty$

$$F_k(u) \longrightarrow 0 \quad \text{as } k \to \infty$$
$$\Rightarrow P(U_k \le u) \to 0$$

$$P(U \le u)$$

As before, we have shown that since

$$U - T \xrightarrow{P} 0 \quad \& \quad T - \sum_{i=1}^{k} D_i \xrightarrow{P} 0$$
$$\Rightarrow U - T \xrightarrow{d} 0 \quad \& \quad T - \sum_{i=1}^{k} D_i \xrightarrow{d} 0$$
$$\Rightarrow (U - T) + (T - \sum_{i=1}^{k} D_i) \xrightarrow{d} 0$$
$$\Rightarrow U - \sum_{i=1}^{k} D_i \xrightarrow{d} 0$$
$$\Rightarrow U \xrightarrow{d} \text{Dist}\left(\lim_{k\to\infty} \sum_{i=1}^{k} D_i\right)$$

$$Y = \sum_{i=1}^{k} D_i \sim \text{Bin}(k, P(x_i = 1))$$
$$\sim \text{Bin}(k, k e^{-k})$$

$$M_Y(t) = (1 - p + pe^t)^k$$
$$= (1 - ke^{-k} + ke^{-k+t})^k$$
$$= \left(1 - \frac{k^2 e^{-k}(1 - e^t)}{k}\right)^k$$

$$\lim_{k\to\infty} \frac{k^2}{e^k} \overset{LH}{=} \lim_{k\to\infty} \frac{2k}{e^k} \overset{LH}{=} \lim_{k\to\infty} \frac{2}{e^k} = 0$$

$$\Rightarrow \lim_{k\to\infty} M_Y(t) = e^0 = 1$$

2). $Y_1, \ldots, Y_n \overset{iid}{\sim} E(Y_i) = \mu, \quad Var(Y_i) = \sigma^2 < \infty$

$R_1, \ldots, R_n = I(Y_i \text{ obs}) \rightarrow R_i = 0, 1$

$x_1, \ldots, x_n \overset{iid}{\sim}$ additional info

R_i, Y_i indep given x_i

(Y_i, R_i, x_i) iid Random vectors.

$\pi(x) = P(R_i = 1 | x_i = x) \rightarrow$ known & bounded by a positive constant from below for any x in the support of x_i

(a) $\hat{\mu}_1 = \dfrac{\sum_{i=1}^n R_i Y_i}{\sum_{i=1}^n R_i}$

(i) Derive the asymptotic limit of $\hat{\mu}_1$ (μ^*) & derive the (ii) asymptotic distribution of $\sqrt{n}(\hat{\mu}_1 - \mu^*)$.

- leave expressions in the result.

(i) $\hat{\mu}_1 = \dfrac{\sum_{i=1}^n R_i Y_i}{n} \cdot \dfrac{1}{\sum_{i=1}^n R_i / n}$

By the strong law of large #s,

$\dfrac{\sum_{i=1}^n R_i}{n} \xrightarrow{a.s.} E[R_i] = E[I(R_i = 1)] = P(R_i = 1) = \pi(x)$

$\dfrac{\sum_{i=1}^n R_i Y_i}{n} \xrightarrow{a.s.} E[R_i Y_i] = E[E[R_i Y_i | x_i]]$

$= E[R_i E[Y_i | x_i]] = E[R_i \mu] = \mu E[R_i]$ } see work in part (b)

$= \mu \pi(x)$

by continuous mapping $\Rightarrow \dfrac{(\sum_{i=1}^n R_i Y_i / n)}{(\sum_{i=1}^n R_i / n)} \xrightarrow{a.s.} \dfrac{\mu \pi(x)}{\pi(x)} = \mu$ ✓

Continued $\longrightarrow$

(ii) $\sqrt{n}(\hat{\mu}_1 - \mu^*)$ asymptotic dist ?

By the multivariate central limit thm,

$$\sqrt{n}\left(\begin{bmatrix}\sum_{i=1}^n R_iY_i/n \\ \sum_{i=1}^n R_i/n\end{bmatrix} - \begin{bmatrix}\mu\pi(x) \\ \pi(x)\end{bmatrix}\right) \xrightarrow{d}$$

$$\xrightarrow{d} N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}Var(R_iY_i) & Cov(R_iY_i, R_i) \\ Cov(R_iY_i, R_i) & Var(R_i)\end{bmatrix}\right)$$

$$Var(R_iY_i) = E[(R_iY_i)^2] - (E[R_iY_i])^2$$
$$= \sigma_{11} \quad \ldots$$

$$Cov(R_iY_i, R_i) = E[R_i^2 Y_i] - E[R_iY_i]\,E[R_i]$$
$$= \sigma_{12} \quad \ldots$$

$$Var(R_i) = \pi(x)(1-\pi(x)) = \sigma_{22}$$

By the Delta Method,

$$\sqrt{n}\left(\frac{\sum_{i=1}^n R_iY_i}{\sum_{i=1}^n R_i} - \mu\right) \xrightarrow{d} N(0, \nabla g\, \Sigma\, \nabla g^T)$$

$$\nabla g = [\partial g/\partial x \quad \partial g/\partial y]$$

$$g = x/y \qquad \left(x = \sum_{i=1}^n R_iY_i/n, \quad y = \sum_{i=1}^n R_i/n\right)$$

$$\nabla g = [1/y, -x/y^2] \text{ evaluated at } \underset{\sim}{\mu}:$$

$$= [1/\pi(x), -\mu/\pi(x)]$$

$$= [1/\pi(x), -\mu/\pi(x)]\begin{bmatrix}\sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22}\end{bmatrix}\begin{bmatrix}1/\pi(x) \\ -\mu/\pi(x)\end{bmatrix}$$

(b) $\hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{R_i Y_i}{\pi(X_i)}$

- Show $\hat{\mu}_2$ consistent for μ & derive asymptotic dist of $\sqrt{n}(\hat{\mu}_2 - \mu)$.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(\frac{R_i Y_i}{\pi(x_i)}\right)$$

By the strong law of large #s,

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{R_i Y_i}{\pi(x_i)}\right) \xrightarrow{a.s.} E\left[\frac{R_i Y_i}{\pi(x_i)}\right] = \frac{1}{\pi(x_i)} E[R_i Y_i]$$

$$= \frac{1}{\pi(x_i)} E[E[R_i Y_i \mid x_i]] = \frac{E[E(R_i \mid x_i] E[Y_i \mid x_i]]}{\pi(x_i)}$$

since given x_i, R_i & Y_i indep

$$= \frac{E[\cancel{\pi(x_i)}\, E[Y_i \mid x_i]]}{\cancel{\pi(x_i)}}$$

$$= E[E[Y_i \mid x_i]] = E[Y_i] = \mu \checkmark$$

$$\sqrt{n}(\hat{\mu}_2 - \mu) \xrightarrow{d} ? \equiv N(0, Var(R_i Y_i / \pi(x_i)))$$

$$Var(R_i Y_i / \pi(x_i)) = \frac{1}{(\pi(x_i))^2} Var(R_i Y_i)$$

$$= \frac{E[(R_i Y_i)^2] - (E[R_i Y_i])^2}{(\pi(x_i))^2}$$

$$= \frac{E[E[R_i^2 \mid x_i] E(Y_i^2 \mid x_i)] - (E[E[R_i \mid x_i] E[Y_i \mid x_i]])^2}{(\pi(x_i))^2}$$

$$= \frac{\pi(x_i) E[E(Y_i^2 \mid x_i]] - (\pi(x_i))^2 \mu^2}{(\pi(x_i))^2}$$

$$= \frac{(\mu + \sigma^2) - \pi(x_i)\mu^2}{\pi(x_i)}$$

$\longrightarrow$

By the Liapunove CLT,

$$\frac{\sum_{i=1}^{n}(x_{ni}-\mu_{ni})}{\sigma_n} \xrightarrow{d} N(0,1)$$

$$\mu_{ni} = E[X_{ni}] = \mu$$

$$\sigma_{ni}^2 = Var(X_{ni}) = \frac{(\mu+\sigma^2) - \pi(x_i)\mu^2}{\pi(x_i)}$$

$$\sigma_n = \sqrt{\sum_{i=1}^{n} \sigma_{ni}^2}$$

$$= \sqrt{\sum_{i=1}^{n} \frac{(\mu+\sigma^2) - \pi(x_i)\mu^2}{\pi(x_i)}}$$

(c) $g(x_i)$ = measurable function w/ finite second moment

$$\hat{\mu}_g = \frac{1}{n}\left[\sum_{i=1}^{n} \frac{R_i Y_i}{\pi(x_i)} + \sum_{i=1}^{n} \frac{(1 - R_i/\pi(x_i))}{g(x_i)}\right]$$

$$= \hat{\mu}_2 + \frac{1}{n}\sum_{i=1}^{n} \frac{(1 - R_i/\pi(x_i))}{g(x_i)}$$

Show $\hat{\mu}_g$ is a consistent estimator for μ
& derive the asymptotic dist of $\sqrt{n}(\hat{\mu}_g - \mu)$

We already know $\hat{\mu}_2 \xrightarrow{a.s.} \mu$.
If we can show $\frac{1}{n}\sum_{i=1}^{n} \frac{(1 - R_i/\pi(x_i))}{g(x_i)} \xrightarrow{a.s.} 0$, then
we have proved $\hat{\mu}_g$ is consistent.

By the strong law of large #s,

$$\frac{1}{n}\sum_{i=1}^{n} \frac{(1 - R_i/\pi(x_i))}{g(x_i)} \xrightarrow{a.s.} \frac{1 - E[R_i/\pi(x_i)]}{g(x_i)}$$

$$= \frac{1 - E[R_i]/\pi(x_i)}{g(x_i)}$$

$$= \frac{1 - \pi(x_i)/\pi(x_i)}{g(x_i)} = \frac{1-1}{g(x_i)}$$

$= 0$ given $g(x_i) \neq 0$

Asymptotic Dist of $\sqrt{n}(\hat{\mu}_g - \mu)$:

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{(1 - R_i/\pi(x_i))}{g(x_i)} - 0\right) \xrightarrow{d} N\left(0, \overbrace{Var\left(\frac{1 - R_i/\pi(x_i)}{g(x_i)}\right)}^{?}\right)$$

3). (a) T_0 = unbiased estimator of θ

squared error loss.

(i) Show $T_0 + c$ is not a minimax estimator under squared error loss.

Minimax estimators are admissible

Is $T_0 + c$ admissible?

Admissible: $R(\theta, d) \le R(\theta, d')$ $\forall \theta$ and

Strict inequality for some θ. (d admissible)

Inadmissible (d): $R(\theta, d) > R(\theta, d')$ for some θ.

$R(\theta, d^*)$ for $d^* = T_0 + c$

$= E[(\theta - d^*)^2]$

$= E[(\theta - T_0 - c)^2]$

$= E[(\theta - T_0)^2 - 2c(\theta - T_0) + c^2]$

$= Var(T_0) - 2c(0) + c^2$

$= Var(T_0) + c^2$

$R(\theta, T_0) = E[(\theta - T_0)^2] = Var(T_0)$

Since $Var(T_0) < Var(T_0) + c^2$

$\Rightarrow$ $R(\theta, T_0) < R(\theta, d^*)$

$\Rightarrow$ $d^* = T_0 + c$ inadmissible

$\Rightarrow$ d^* <u>not</u> minimax ✓

If $R(\theta, T_0) < R(\theta, T_0 + c)$ $\forall \theta$ for $c \neq 0$

$\Rightarrow$ $\sup_\theta R(\theta, T_0) < \sup_\theta R(\theta, T_0 + c)$

$\Rightarrow$ $T_0 + c$ not minimax

(a) (ii) Show that the estimator cT_0 is <u>not</u> minimax under squared error loss unless $\sup_\theta R_T(\theta) = \infty$ for any estimator T of θ $(c \in (0,1))$

$$R(\theta, cT_0) = E[(\theta - cT_0)^2]$$
$$= E[(\theta - T_0 + T_0 - cT_0)^2]$$
$$= E[(\theta - T_0)^2 + 2(\theta - T_0)(T_0 - cT_0) + T_0^2(1-c)^2]$$
$$= Var(T_0) + 0 + (1-c)^2 E[T_0^2]$$
$$= Var(T_0) + (1-c)^2 (Var(T_0) + (E(T_0))^2]$$
$$= Var(T_0) + (1-c)^2 (Var(T_0) + \theta^2)$$

$$R(\theta, T) = E[(T-\theta)^2]$$

T = estimator of θ → may be unbiased or not

let $E[T] = g(\theta)$

$$= E[(T - g(\theta) + g(\theta) - \theta)^2]$$
$$= E[(T - g(\theta))^2 + 2(T - g(\theta))(g(\theta) - \theta) + (g(\theta) - \theta)^2]$$
$$= Var(T) + 0 + (g(\theta) - \theta)^2$$

$\sup_\theta R(\theta, cT_0) \leq \sup_\theta R(\theta, T) \quad \forall \theta$ if

unbiased T_0

$$\sup_\theta \left[Var(T_0) + (1-c)^2 (Var(T_0) + \theta^2)\right] \leq$$

$$\sup_\theta \left[Var(T) + (g(\theta) - \theta)^2 \right]$$