Stats 510 Assignment 5.

19342932 Shu Zhou

1. $F(x) = P(X \le x, W=1) + P(X \le x, W=-1)$

$= P(WT \le x, W=1) + P(WT \le x, W=-1)$

Hence, we can summarize $F(x)$ into $F(x) = \frac{2}{3}P(T \le x) + \frac{1}{3}P(-T \le x)$

For $x<0$, $P(-T \le x) = P(T \ge -x) = e^{\lambda x} \Rightarrow F(x) = \frac{1}{3}e^{\lambda x}$. $(x<0)$

Hence $f(x) = F'(x) = \frac{1}{3}\lambda e^{\lambda x}$ $(x<0)$

For $x>0$. $F(x) = 1 - P(T > x) = 1 - \frac{2}{3}e^{-\lambda x}$

Hence $f(x) = \frac{2}{3}\lambda e^{-\lambda x}$ $(x>0)$

Hence, in total $f(x) = \begin{cases} \frac{1}{3}\lambda e^{\lambda x} & (x<0) \\ \frac{2}{3}\lambda e^{-\lambda x} & (x>0) \end{cases}$

2. (a). $f(t) = f(-t)$; it is obviously that $f(x) = f(-x)$ for all x, i.e. f is evenly distributed. Hence, X and $-X$ have the same distribution, i.e X is distributed symmetrically.

On the other hand, if X and $-X$ have the same distribution; suppose $G(x)$ is the distribution of $-X$

$G(x) = P(-X \le x) = 1 - P(X \le -x) = 1 - F(-x) = F(x)$.

Then $g(x) = G'(x) = \frac{d}{dx}(1 - F(-x)) = f(-x)$.

Hence. Q.E.D.

(b) $\int_{-\infty}^{\infty} |x| f(x) dx = \int_0^{\infty} x f(x) dx + \int_{-\infty}^{0} -x f(x) dx$

According to (a), we know that f is evenly distributed.

Hence $\int_0^{\infty} x f(x) dx + \int_{-\infty}^{0} -x f(x) dx$

$= \int_0^{\infty} x f(x) dx + \int_0^{\infty} -x f(x) dx = \int_0^{\infty} x f(x) dx - \int_0^{\infty} x f(x) dx = 0$.

Hence. $E(x) = 0$.

(c.) $\bar{V} = (2B-1)V$

Then for B: $B = \begin{cases} 0 & \text{with prob. } \theta \\ 1 & \text{with prob. } 1-\theta. \end{cases}$

Then $P(\bar{V} \le \bar{u}) = P((2B-1)V \le \bar{u}) = P(B=1, V \le \bar{u}) + P(B=0, -V \le \bar{u})$; Since B and V are independent.

$= P(B=1) \cdot P(V \le \bar{u}) + P(B=0) \cdot P(-V \le \bar{u})$

$= \theta P(V \le \bar{u}) + (1-\theta) P(-V \le \bar{u})$

$= P(V \le \bar{u})$ = p.m.f of V. Since the p.m.f is the same

Hence the density function of the random variable $\bar{V}$ is the same as the density function of V.

3 $P(W=m) = \sum_{i=0} P(P_1 = i, P_2 = m-i)$

(a) Since P_1 and P_2 are independent; Hence

$$P(W=m) = \sum_{i=0}^{m} P(P_1 = i) P(P_2 = m-i)$$
$$= \sum_{i=0}^{m} \frac{e^{-\lambda_1} \lambda_1^{i}}{i!} \frac{e^{-\lambda_2} \lambda_2^{m-i}}{(m-i)!}$$
$$= \sum_{i=0}^{m} \frac{i!}{i!(m-i)!} \times \frac{e^{-\lambda_1} e^{-\lambda_2} \lambda_1^{i} \lambda_2^{m-i}}{i!}$$
$$= \frac{e^{-\theta}}{m!} \left(\sum_{i=0}^{m} \binom{m}{i} \lambda_1^{i} \lambda_2^{m-i} \right)$$ Since the part in the bracket follows the binomial distribution

Hence $= \frac{e^{-\theta}}{m!} (\lambda_1 + \lambda_2)^m$

$$= \frac{e^{-\theta} \theta^m}{m!} \quad (Q.E.D)$$

(b) Since $X_1 + X_2$ follows the possion distribution with parameter $\lambda_1 + \lambda_2$.

We let $X_{1,2} = X_1 + X_2$

$\lambda_{1,2} = \lambda_1 + \lambda_2$

Hence for $X_{1,2} + X_3$, it follows the possion distribution with parameter $\lambda_1 + \lambda_2 + \lambda_3$.

Hence, according to mathematical induction, $X_1 + X_2 + \cdots X_n$ is a Possion random variable, with parameter $\lambda_1 + \lambda_2 + \cdots + \lambda_n$

(c) $\tilde{S}_n = (S_n - \Lambda_n)/\sqrt{\Lambda_n}$

$P(\tilde{S}_n = m) = \frac{e^{-\Lambda_n} \Lambda_n^m - \Lambda_n}{\sqrt{\Lambda_n} \times m!}$, Then the moment generating function.

$$MGF = E[e^{tx}] = \sum_{x=0}^{\infty} p(x) e^{tx}$$
$$= \sum_{x=0}^{\infty} \frac{e^{-\Lambda_n} \Lambda_n^{x} - \Lambda_n}{\sqrt{\Lambda_n} \times x!} \cdot e^{tx}$$
$$= \sum_{x=0}^{\infty} \frac{e^{-\Lambda_n} \Lambda_n^{x}}{\sqrt{\Lambda_n} \times x!} e^{tx} - \frac{\Lambda_n e^{tx}}{\sqrt{\Lambda_n} \times x!}$$
$$= \frac{e^{-\Lambda_n}}{\sqrt{\Lambda_n}} \sum_{x=0}^{\infty} \frac{(\Lambda_n e^{t})^{x}}{x!} - \frac{\Lambda_n}{\sqrt{\Lambda_n}} \sum_{x=0}^{\infty} \frac{e^{tx}}{x!} \quad \text{Since } \sum_{x=0}^{\infty} \frac{e^{tx}}{x!} = e^{e^t}$$
$$= e^{-t\sqrt{\Lambda_n} + \Lambda_n (e^{t/\sqrt{\Lambda_n}} - 1)}$$

Hence we take the logarithm, it gives

$$\log(MGF(t)) = \Lambda_n \left[e^{t/\sqrt{\Lambda_n}} - \frac{t}{\sqrt{\Lambda_n}} - 1 \right]$$

Hence $\lim_{n \to \infty} [\log MGF(t)] = \frac{t^2}{2}$

Hence $\lim_{n \to \infty} MGF(t) = e^{\frac{t^2}{2}}$; Hence the limit distribution is $e^{\frac{t^2}{2}}$

4. first we calculate $P[Y=y]$ directly.

$P[Y=y] = P(y \text{ failure before the } r^{th} \text{ success})$

$= \binom{r+y-1}{y} p^r (1-p)^y$

Hence the MGF of Y is

$M_Y(t) = E(e^{tY}) = \sum_{y=0}^{\infty} \binom{r+y-1}{y} e^{ty} p^r (1-p)^y = \frac{p^r}{(1-e^t(1-p))^r}$

Hence when $r \to \infty$ and $p \to 1$, $r(1-p) \to \lambda$ for some $\lambda \to 0$

$E(Y^2) = \frac{r(1-p)p + r(r+1)(1-p)^2}{p^2} = \frac{\lambda + \lambda^2}{1^2} = \lambda + \lambda^2$

$E(Y) = \frac{r(1-p)}{p} = \lambda$

Hence $Var(Y) = E(Y^2) - [E(Y)]^2 = \lambda$, according to the central limit theorem.

$P(Y=y) \to e^{-\lambda}\frac{\lambda^y}{y!}$; Hence Q.E.D.

5. $p_D = N_D/N$; According to our definition

(a) $x_j = \begin{cases} 1 & \text{(if the } j^{th} \text{ individual is D)} \\ 0 & \text{(or not)} \end{cases}$

Among all possible permutations of size n from a population of N distinct variables

The number of permutations are D's in the position i_j is $\binom{N_D-1}{N-1}$

Hence $P(x_{i1}=1) = \frac{\binom{N_D-1}{N-1}}{\binom{N}{n}}$ $\Rightarrow P(x_{i1}=\varepsilon_1) = \frac{\binom{N_D-\varepsilon_1}{N-\varepsilon_1}}{\binom{N}{n}}$;

$P(x_{i1}=0) = 1 - \frac{\binom{N_D-1}{N-1}}{\binom{N}{n}}$

Then we consider $P(x_{i1}=\varepsilon_1, P x_{i2}=\varepsilon_2)$ number of permutation are D's in position i_2.

$P(x_{i1}=\varepsilon_1, x_{i2}=1) = \frac{\binom{N_D-\varepsilon_1-1}{N-\varepsilon_1-1}}{\binom{N}{n}}$ $\Rightarrow$ Hence, according to mathematical induction

$P(x_{i1}=\varepsilon_1, x_{i2}=\varepsilon_2, \cdots x_{in}=\varepsilon_n) = \frac{\binom{N_D-\sum\varepsilon_i}{N-\sum\varepsilon_i}}{\binom{N}{n}}$

(b) $P(S_n=m) = \frac{\binom{N_D}{m}\binom{N-N_D}{n-m}}{\binom{N}{n}}$ (According to hypergeometric distribution)

(c) In the case of $n=N$.

S_n/n is the average value of the first n of x_i's

$E(S_n/n) = \frac{\sum_{i=1}^{n} E(x_i)}{n}$; where $E(x_i) = \frac{N_d}{N}$

Hence $E(S_n/n) = \frac{N_d}{N}$

$Var(S_n/n) = 0$; since we take the whole set as the sample, and there should be no variance.

6. $X \sim f(x-\theta_0)$; $f(0) > f(x)$ for all $x \neq 0$, $f(x) = f(-x)$ ①

Hence, If we want to maximize the interval $P(X \in I_h)$, we need to maximize $P(x \in [\lambda, \lambda+2h])$

I.e find the value λ, that maximizes $F_X(\lambda+2h) - F_X(\lambda)$.

$\frac{d}{dx}(F_X(\lambda+2h) - F_X(\lambda)) = f_X(\lambda+2h) - f_X(\lambda) = 0 \Rightarrow f_X(\lambda+2h) = f_X(\lambda)$, according to ①

Hence $\lambda+2h$ and λ should be centered at $0 \Rightarrow$ i.e. $\lambda+h=0 \Rightarrow \lambda=-h$.

Hence, the required interval $(x-\theta_0=-h, x-\theta_0=h) \Rightarrow x \in (\theta_0-h, \theta_0+h)$