1 (a) The "Poisson distribution" on the sample space {0,1,2,3,4, ...} assigns probability

$$\frac{1}{k!} \lambda^k e^{-\lambda}$$

to the set {k} for each k=0,1,2,3,4...
Show that this is a probability measure on the set {0,1,2, } (i.e. show the total measure =1)
(b) Explain how the Poisson distribution arises as the model for the number of occurrences in time 1 of an event which has the property that the limit

limit (1/t) (number of occurrences in every interval of length t) = λ
$t \to 0, t>0$

(Suggestion: Subdivide the interval of length 1 into N pieces, N large positive integer, N so large that probability of two occurrences in one interval of length 1/N can be neglected).

(a) $$\sum_{k=1}^{+\infty} \frac{1}{k!} \lambda^k e^{-\lambda} = 1$$

$$e^{-\lambda} \underbrace{\sum_{k=1}^{+\infty} \frac{1}{k!} \lambda^k}_{\text{Taylor series for } e^{\lambda}} = 1$$

$$e^{-\lambda} \cdot e^{\lambda} = 1$$

$$1 = 1 \checkmark \quad \boxed{QED}$$

$$P(k \text{ occurences}) = \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}$$

$$\lim_{n\to+\infty} P(k \text{ occurences}) = \frac{\lambda^k}{k!} \lim_{n\to+\infty} \frac{(n)\cdots(n-k+1)}{n^k} \cdot \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}$$

$$P(k \text{ occurences}) = \frac{\lambda^k}{k!} \lim_{n\to+\infty} \left(1-\frac{\lambda}{n}\right)^n$$

$$\boxed{P(k \text{ occurences}) = \frac{\lambda^k}{k!} e^{-\lambda}}$$

$\boxed{QED}$

(b)

Subdivide interval of length 1 into n subintervals

let $\lambda = \lim_{t\to 0} \frac{1}{t}$ (number of occurences in every interval of length t)

Neglecting probability of >1 occurence in one subinterval, P of k occurrences in interval of length 1 with n subintervals $= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$

2 A fair coin is tossed 10,000 times. Use the approximation of binomial by normal to find

(a) the probability that the number of heads is less than 5,075

(b) the smallest number h such that the probability of no more than h heads is [redacted] $\geq .99$

(normal distribution table is attached)

(a) $\mu = np = 10{,}000\left(\frac{1}{2}\right) = 5{,}000$

$\sigma^2 = np(1-p) = 10{,}000\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 2{,}500 \Rightarrow \sigma = \sqrt{2500} = 50$

$z = \dfrac{5075 - 5000}{50} = \dfrac{75}{50} = 1.5$

$\Phi(z = 1.5) = \boxed{0.9332}$

(b) smallest z such that $\Phi(z) \geq 0.99$

(from table) $z = 2.303$

$2.303 = \dfrac{h - 5{,}000}{50}$

$115.15 = h - 5{,}000$

$h = 5115.15 \approx 5116$

$\boxed{h = 5116}$

$$\begin{array}{r} 2.303 \\ \times \quad 50 \\ \hline 0000 \\ +\ 11515 \\ \hline 115.150 \end{array}$$

3 (a) Explain why a non-negative random variable Y with expected value E satisfies: probability measure of {x in sample space: Y(x) >a} $\leqslant$ E/a (for each a>0).
(b) Use part (a) to show that, if X is a random variable with mean 0 and finite standard deviation and if X_1, X_2, X_3, X_4 ... are independent random variables with the same distribution as X, then for each e>0

$$\lim_{n\to+\infty} \text{probability measure of } \{ x \text{ in sample space} : |(1/n)|X_1(x)+X_2(x)+\dots X_n(x)|>e\} = 0$$

(a) $E(Y) = \int_{-\infty}^{+\infty} t f_Y(t)\,dt = \int_{0}^{+\infty} t f_Y(t)\,dt$

because nonnegative random variable

$$\int_0^{+\infty} t f_Y(t)dt = \int_0^a t f_Y(t)dt + \int_a^{+\infty} t f_Y(t)\,dt \geq \int_a^{+\infty} t f_Y(t)\,dt$$

$$\int_a^{+\infty} t f_Y(t)dt \geq \int_a^{+\infty} a f_Y(t)dt = a\int_a^{+\infty} f_Y(t)dt = a P(Y \geq a)$$

Therefore,

$$E(Y) \geq a P(Y \geq a)$$

$$\boxed{P(Y \geq a) \leq \frac{E(Y)}{a}}$$ QED

(b) let $S_n = \dfrac{X_1 + \dots + X_n}{n}$

$E(S_n) = \frac{1}{n}[\underbrace{E(X_1) + \dots + E(X_n)}_{=0}]$ (since $X_1, \dots, X_n$ are IID)

$Var(S_n) = \frac{1}{n^2}[Var(X_1) + \dots + Var(X_n)]$ (since $X_1, \dots, X_n$ are IID)

$= \frac{n\,Var(X)}{n^2} = \frac{Var\,X}{n}$

$\forall \varepsilon > 0,$

$$P(|S_n - E(S_n)| \geq \varepsilon) = P(|S_n| \geq \varepsilon) = P(S_n^2 \geq \varepsilon^2) \leq \frac{E((S_n - E(S_n))^2)}{\varepsilon^2}$$

$$P(|S_n| \geq \varepsilon) \leq \frac{Var(S_n)}{\varepsilon^2} = \frac{Var(X)}{n\varepsilon^2}$$

since $0 \leq P(|S_n| \geq \varepsilon)$,

$$\lim_{n\to+\infty} P(|S_n| \geq \varepsilon) \leq \lim_{n\to+\infty} \frac{Var(X)}{n\varepsilon^2} = 0, \therefore \boxed{\lim_{n\to+\infty} P(|S_n| \geq \varepsilon) = 0}$$ QED

4 Suppose $X_1, X_2, X_3, \ldots$ are independent identically distributed random variables with mean=0 and and $E(X_1^4) < +\infty$. Show that

$$\sum_{n=1}^{+\infty} E[(1/n^4)(X_1+X_2+X_3 \ldots + X_n)^4] < +\infty$$

[This is part of the proof of the Strong Law of Large Numbers]

$$\sum_{n=1}^{+\infty} \frac{1}{n^4} E[(X_1 + X_2 + \ldots + X_n)^4]$$

$$= \sum_{n=1}^{+\infty} \frac{1}{n^4} \sum_{a=1}^{n}\sum_{b=1}^{n}\sum_{c=1}^{n}\sum_{d=1}^{n} E(X_a X_b X_c X_d)$$

since $E(X) = 0$ and $X_1, \ldots, X_n$ are independent,

if $a \neq b \neq c \neq d$, then $E(X_a X_b X_c X_d) = \cancel{E(X_a)}^{0}\cancel{E(X_b)}^{0}\cancel{E(X_c)}^{0}\cancel{E(X_d)}^{0} = 0$

and if one index doesn't equal the other 3 (for example, $a \neq b = c = d$), then $E(X_a X_b X_c X_d) = \cancel{E(X_a)}^{0} E(X_b X_c X_d) = 0$

Thus, there are only terms of the form $E(X_i^4)$ and $E(X_i^2 X_j^2)$ left in the summation. There are n terms of the form $E(X_i^4)$ and $3n(n-1)$ terms of the form $E(X_i^2 X_j^2)$.

So,

$$= \sum_{n=1}^{+\infty} \frac{1}{n^4} \left(n E(X_i^4) + 3n(n-1) E(X_i^2 X_j^2)\right)$$

$a = b \neq c = d$
$a = c \neq b = d$
$a = d \neq b = c$ } 3

n ways to pick $n-1$ ways to pick

$E(X_i^2 X_j^2) \leq E(X_i^4)$

(since $xy \leq \frac{x^2+y^2}{2}$),

so

$$n E(X_i^4) + 3n(n-1) E(X_i^2 X_j^2) \leq (n + 3n(n-1)) E(X_i^4)$$
$$\leq 3n^2 E(X_i^4)$$

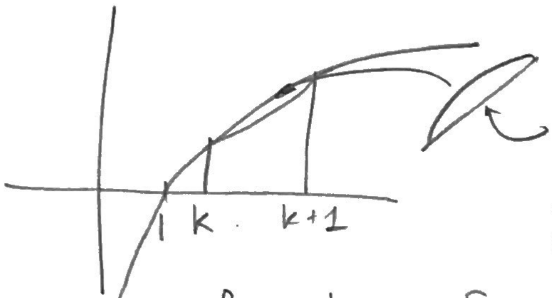
Thus,

$$\leq \sum_{n=1}^{+\infty} \frac{1}{n^4}\left(3n^2 E(X_i^4)\right) = \sum_{n=1}^{+\infty} \frac{3}{n^2} E(X_i^4) = E(X_i^4) \sum_{n=1}^{+\infty} \frac{3}{n^2} < +\infty$$

converging series

In summary, $\boxed{\sum_{n=1}^{+\infty} \frac{1}{n^4} E[(X_1 + X_2 + \ldots + X_n)^4] < +\infty}$ $\boxed{QED}$

5 Use the trapezoid rule for the integral of ln x to establish Stirling's Formula up to a constant factor.



$$m = \frac{\ln(k+1) - \ln k}{k+1-k} = \ln(k+1) - \ln k$$

$$y - \ln k = [\ln(k+1) - \ln(k)](x-k)$$

$$f_k = \ln x - [[\ln(k+1) - \ln(k)](x-k) + \ln k]$$

(*) $|f_k| \le \max|f_k''|$

$|f_k| = f_k$ $\qquad f_k' = \frac{1}{x} - (\ln(k+1) - \ln(k))$

$$f_k'' = -\frac{1}{x^2}$$

$$|f_k''| = \frac{1}{x^2} \Rightarrow \max|f_k''| = \frac{1}{k^2}$$

So, $|f_k| \le \frac{1}{k^2}$

$$\int_k^{k+1} f_k\,dx \le \int_k^{k+1} |f_k|\,dx \le \int_k^{k+1} \frac{1}{k^2}\,dx = \frac{1}{k^2} \Rightarrow \int_k^{k+1} f_k\,dx \le \frac{1}{k^2}$$

let $E_n = \sum_{k=1}^{n-1} \int_k^{k+1} f_k\,dx \le \sum_{k=1}^{n-1} \frac{1}{k^2}$. Then, $\lim_{n\to+\infty} E_n \le \underbrace{\sum_{k=1}^{+\infty} \frac{1}{k^2}}_{\text{converging series}} = c$

a constant c

trapezoidal area approximation

$$E_n = \int_1^n \ln x\,dx - \Big(\underbrace{\cancel{\tfrac{1}{2}\ln 1} + \tfrac{1}{2}\ln 2}_{\text{area of first trapezoid}} + \tfrac{1}{2}\ln 2 + \tfrac{1}{2}\ln 3 + \dots + \tfrac{1}{2}\ln(n-1) + \tfrac{1}{2}\ln n\Big)$$

summing up areas of trapezoids of width 1

$$E_n = (x\ln x - x)\Big|_1^n - \left[\ln(n!) - \tfrac{1}{2}\ln(n)\right]$$

integration by parts

sum of logs is log of product

$$E_n = n\ln(n) - n - [1\ln(1) - 1] - \ln(n!) + \tfrac{1}{2}\ln(n)$$

$$E_n = n\ln(n) - n + 1 - \ln(n!) + \tfrac{1}{2}\ln(n)$$

(continued on back)

$$\lim_{n\to+\infty} E_n = c = \lim_{n\to+\infty} n\ln(n) - n + 1 - \ln(n!) + \tfrac{1}{2}\ln(n)$$

$e \quad e \quad e$ (exponentiate)

$$e^c = \lim_{n\to+\infty} e^{n\ln(n) - n + 1 - \ln(n!) + \frac{1}{2}\ln(n)}$$

$$C = \lim_{n\to+\infty} \left(e^{\ln(n)}\right)^n \cdot e^{-n} \cdot e^{1} \cdot \left(e^{\ln(n!)}\right)^{-1} \cdot \left(e^{\ln(n)}\right)^{1/2}$$

$$\frac{C}{e} = \lim_{n\to+\infty} \frac{n^n \cdot \sqrt{n}}{e^n \cdot n!}$$

$$\boxed{\lim_{n\to+\infty} \frac{\left(\frac{n}{e}\right)^n \cdot \sqrt{n}}{n!} = C}$$

$\boxed{\text{QED}}$