Note Title 1/31/2011

Let $k' = k-1$, $n' = n-1$

$$= \sum_{k'=0}^{n'} n \cdot \frac{n'!}{(n'-k')!\,k'} p \cdot p^{k'} * (1-p)^{n'-k'}$$

$$= np \left(\underbrace{\sum_{k'=0}^{n'} \frac{n'!}{(n'-k')!\,k'!} p^{k'} (1-p)^{n'-k'}}_{=1} \right)$$

It is the total prob of a binomial R.V w. para. p, n'.

$Var(X) = np(1-p)$ ← will come back later.

* Usually (but not necessarily), a binomial distri models the scenario of tossing a bent coin n times & count the # of heads.

Q: The score of a baseball team is a binomial R.V with para p, n

$E(X) = ?$ $Var(X) = ?$

3. Geometric Random Variable with 1 parameter $0 < p < 1$

$S_X = \{0, 1, \ldots\ldots\}$ all non-negative integers

$p_k = p(1-p)^k$

Example: $p = \frac{1}{4}$. Geometric R.V has

$p_0 = \frac{1}{4}(1-\frac{1}{4})^0$

$p_1 = \frac{1}{4}(1-\frac{1}{4})^1$ . . . . .

* The geometric RV. generally models the number of trials before the first head when flipping an unfair coin with head prob p.

Q The number of cars in a parking lot is a geometric R.V w. para p.

$E(X) = ?$

Ans: $E(X) = \sum_{k=0}^{\infty} p_k \times k$

$$= \sum_{k=0}^{\infty} p(1-p)^k \times k$$

$$= \sum_{k=1}^{\infty} p(1-p)^k \times k$$

$$= \frac{\text{the first term}}{(1-\text{ratio})^2}$$

$$= \frac{p(1-p)}{(1-(1-p))^2} = \frac{1-p}{p}$$

$Var(X) = \frac{1-p}{p^2}$

*

④ Poisson Random Var with para $\alpha > 0$

063

$S = \{0, 1, 2, \dots\}$: all non-negative integers

$p_k = \frac{\alpha^k}{k!} e^{-\alpha}$

$\boxed{\text{We use the convention } 0! = 1.}$

Ex: if $\alpha = 0.5$

$p_0 = \frac{0.5^0}{0!} e^{-0.5} = \frac{1}{1} e^{-0.5} = e^{-0.5}$

$p_1 = \frac{0.5^1}{1!} e^{-0.5} = \frac{0.5}{1} e^{-0.5} = 0.5 e^{-0.5}$

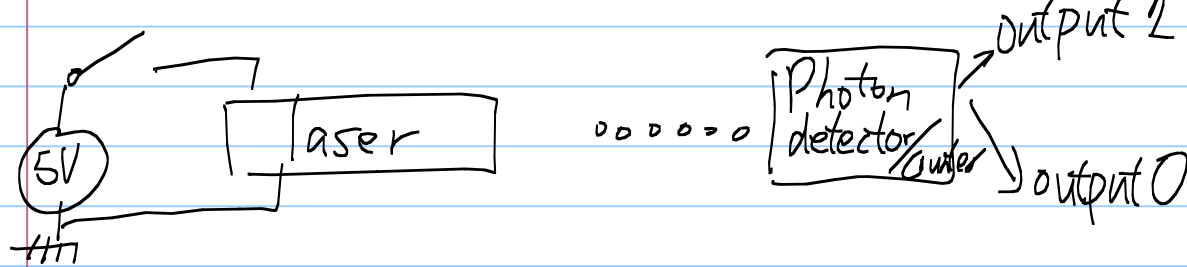
$p_2 = \frac{0.5^2}{2!} e^{-0.5} = \frac{0.5^2}{2} e^{-0.5} = \frac{1}{8} e^{-0.5}$

* The sample space of a Poisson R.V is exactly the same as that of a geometric R.V.

* Poisson random variable is used to model the experiment that ① Consider a fixed time interval, ② "A customer may show up or not" has the same prob for any time instant. ③ Knowing the average number of customers within this interval/duration is α ④ The actual number of customers is a

Poisson R.V w. para α.

* Poisson is quite common, especially in physics.



A laser that can be turned on (5V) or off (0V)

Once it is on, in average 1000 photons/msec.

A photon detector count the number of photons X in 0.1 msec. And output 0 or 1 depending on X.

If there is no "ambient noise"

$$\text{Output} = \begin{cases} 1 & \text{if } X > 0 \\ 0 & \text{otherwise} \end{cases}$$

(but in reality, we choose $\boxed{X > 30}$ instead.

What is the prob that the output

is 0 even if the laser is ON.

Ans: X is a Poisson R.V.

with α = the averge # of photons in 0.1 msec

$= 1000 \times 0.1 = 100$

Then the answer is simple

$$P(X \leq 30) = \sum_{k=0}^{30} \frac{100^k}{k!} \times e^{-100}$$

$$\approx 2 \times 10^{-16}$$

Q: Suppose we reduce the interval to 0.01 msec

Y is the # of photons in 0.01 msec.

$$\text{Output} = \begin{cases} 1 & \text{if } Y \geq 3 \\ 0 & \text{otherwise.} \end{cases}$$

$P(\text{Output} = 0$ given the laser is ON$) = ?$

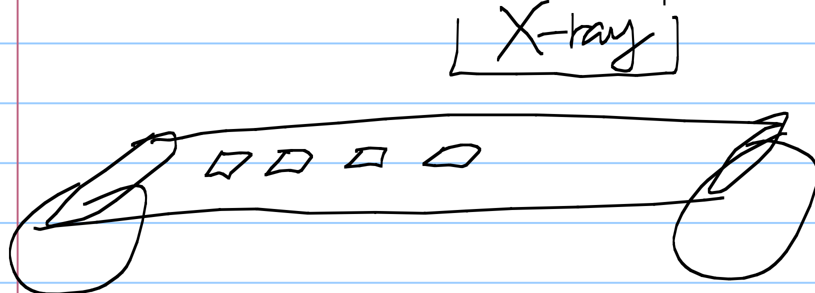
Ans: $\alpha = 1000 \times 0.01 = 10$

$$P(Y \leq 3) = \sum_{k=0}^{3} \frac{10^k}{k!} e^{-10} = 1\%$$

Namely: sending at rate $\frac{1}{0.1 \text{ msec}} = 10k$. the error prob $\approx 2 \times 10^{-16}$ if we increase the rate to 100k.

the bit error rate increases to 1%.
a trade-off between communication speed and the error rate.

Ex: A factory uses X-ray to test the defective chips sequentially



We know that <u>in average</u> we will find 3 defective chips every 20 minutes.

Let X be the number of defective chips found from 1-2:30 pm

Q: $P(X \le 6) = ?$

Ans: Construct the W.A first.

X must be a Poisson

since within 1.5 hours, the avg # of defective chips is $90 \times \frac{3}{20} = \frac{27}{2} = \alpha$

X is Poisson with $\alpha = \frac{27}{2}$

$$P(X \le 6) = \sum_{k=0}^{6} P_k$$

$$= \sum_{k=0}^{6} \frac{\left(\frac{27}{2}\right)^k}{k!} e^{-\frac{27}{2}}, \quad 7 \text{ terms}$$

$= 0.01925362$

Q: $E(X) = ?$

Ans: $$E(X) = \sum_{k=0}^{\infty} P_k \, k$$

$$= \sum_{k=0}^{\infty} \frac{\left(\frac{27}{2}\right)^k}{k!} e^{-\frac{27}{2}} \times k$$

$$= \sum_{k=1}^{\infty} \frac{\left(\frac{27}{2}\right)^k}{(k-1)!} e^{-\frac{27}{2}}$$

Let $k' = k-1$

$$= \left(\frac{27}{2}\right) \underbrace{\sum_{k'=0}^{\infty} \frac{\left(\frac{27}{2}\right)^{k'}}{(k')!} e^{-\frac{27}{2}}}_{\text{Total prob} = 1}$$

$$= \frac{27}{2} = \alpha$$

Q: $E(X(X-1)) = ?$

Ans: $$\sum_{k=0}^{\infty} \frac{\alpha^k}{k!} e^{-\alpha} \left(k \cdot (k-1)\right)$$

$$= \sum_{k=2}^{\infty} \frac{\alpha^k}{(k-2)!} e^{-\alpha}$$

$$= \alpha^2 \sum_{k=2}^{\infty} \frac{\alpha^{k-2}}{(k-2)!} e^{-\alpha}$$

let $k' = k-2$

$$= \alpha^2 \underbrace{\sum_{k'=0}^{\infty} \frac{\alpha^{k'}}{k'} e^{-\alpha}}_{=1}$$

$$= \alpha^2$$

Q: $E(X^2) = ?$ $\quad Var(X) = ?$

Ans: $X^2 = X(X-1) + X$

$$\Rightarrow E(X^2) = E(X(X-1)) + E(X)$$

$$= \alpha^2 + \alpha$$

$$Var(X) = E(X^2) - m^2 = \alpha^2 + \alpha - (\alpha)^2$$

$$= \alpha$$

HW5 Q8

The number of page requests that arrive at a Web server is Poisson w. avg 6000 requests per minute.

Q: P(No request in 100 ms)

Q: P(5 to 10 requests in 100 ms)

Q: If more than 15 requests in 100 ms. The server crashes
P(server crashes) = ?

avg # of requests / 100 ms = 10

Ans: X is the number of packets in 100 ms.

X is Poisson with

$$\alpha = 6000 \times \frac{1}{60} \times 0.1 = 10 \text{ (reques per 100 msec)}$$

$$P(X=0) = \frac{\alpha^0}{0!} e^{-\alpha} = e^{-10} \approx 4.54 \times 10^{-5}$$

A: $$P(5 \leq X \leq 10) = \sum_{k=5}^{10} \frac{10^k}{k!} e^{-10}$$

$$= 55.4\%$$

A: $$P(\text{crashes}) = P(X > 15) = 1 - P(X \leq 15)$$

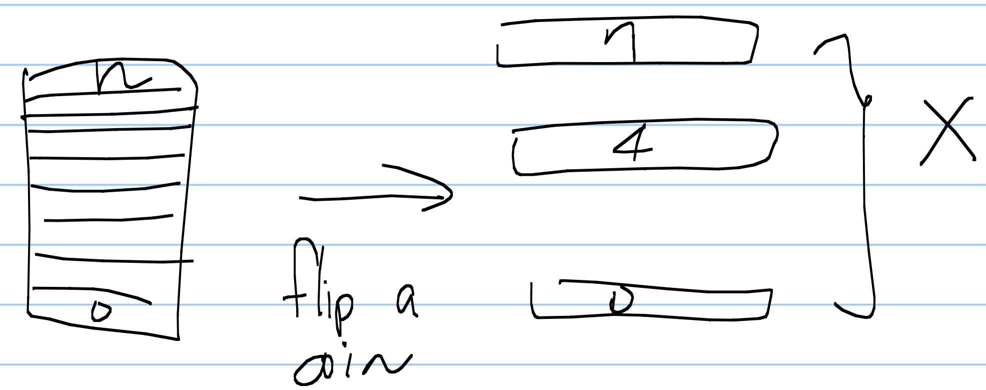
$$= 1 - \sum_{k=0}^{15} \frac{10^k}{k!} e^{-10}$$

$$\approx 4.87\%$$

* The connection between binomial & Poisson distributions.

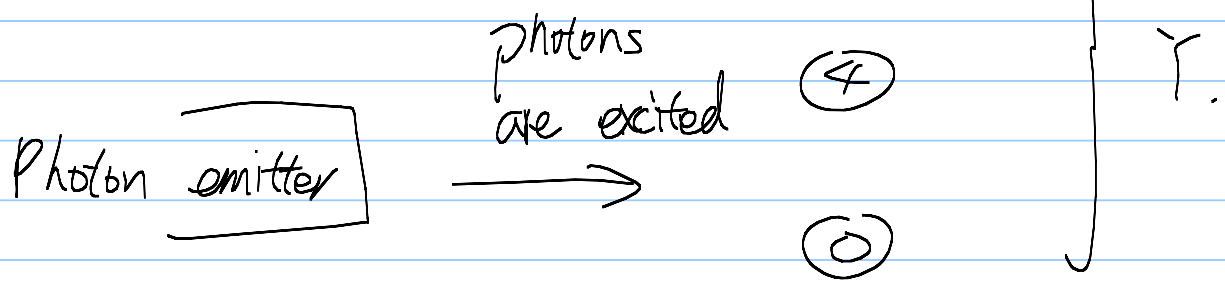
$E(X) = np$

Binomial



$E(Y) = \alpha$

Poisson



The difference is that there are thousands of millions different photon that may go through the laser, but only a very small fraction of them can go through.

I.e. For binomial distribution, we keep $E(X) = np = \alpha$. & let $n \to \infty$, $p = \frac{\alpha}{n}$ so that $E(X) = E(Y)$

Note Title 2/7/2011

Then we have

Binomial

$$P_k = \frac{n!}{k!\,(n-k)!} \cdot p^k (1-p)^{n-k}$$

$$= \frac{\overbrace{n \cdot (n-1) \cdot \cdots \cdot (n-k+1)}^{k \text{ of them}}}{k!} \cdot p^k \cdot (1-p)^n$$

$$\cdot \frac{1}{(1-p)^k}$$

$$\approx \frac{1}{k!} \cdot (np)^k \cdot \left(1 - \frac{\alpha}{n}\right)^n \cdot \frac{1}{1}$$

$$= \frac{\alpha^k}{k!} e^{-\alpha} = P_k \quad \text{Poisson.}$$

* In sum: Poisson is the limit of a binomial with $n \to \infty$, $p = \frac{\alpha}{n}$

1. Many different R.Vs. (discrete thus far)
2. The W.A.
3. Expectation & variance
4. New computation skills.
5. The same counting principle

Continuous R.Vs.

1. Sample space is continuous.
2. The W.A is specified by the prob density function $f_X(x)$

$$P(a \le X \le b) = \int_a^b f_X(x)\,dx$$

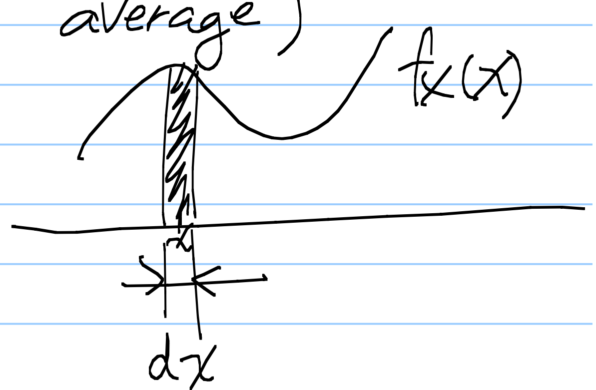
the area underneath $f_X(x)$.

Plotting $f_X(x)$ is just like plotting any function except that $f_X(x) \ge 0$.

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.$$

3. Expectation (Weighted average)

$$E(X) = \int_{-\infty}^{\infty} \underbrace{x}_{\text{face value}} \underbrace{f_X(x)\,dx}_{\text{weight}}$$



$$E(X^2) = \int_{-\infty}^{\infty} \underbrace{x^2}_{\text{face value}} \underbrace{f_X(x)\,dx}_{\text{weight}}$$

$$E(e^X) = \int_{-\infty}^{\infty} e^x f_X(x)\,dx$$

Ex: $f_X(x) = \begin{cases} \frac{1}{3} & \text{if } 0 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$

Q: Find $E(X) = ?$ Ans: $\int x f_X(x) = \int_0^3 x \frac{1}{3}\,dx = \frac{9}{6}$

* Expectation of a constant is the constant itself.

$$E(\pi) = \int_{-\infty}^{\infty} \pi f_X(x)\,dx = \pi \int_{-\infty}^{\infty} f_X(x)\,dx = \pi$$

* Expectation is linear

$$E(a\,g(X)) = a\,E(g(X))$$

$$\because \int_{-\infty}^{\infty} a\,g(x) f_X(x)\,dx = a \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$$

$$E(g_1(X) + g_2(X)) = E(g_1(X)) + E(g_2(X))$$

Again, we use the same formulas of expectation to define the "variance"

$$Var(X) \overset{①}{=} \boxed{E((X-m)^2)} = \int_{-\infty}^{\infty} (x-m)^2 f_X(x)\,dx$$

$$\overset{②}{=} \boxed{E(X^2) - m^2}$$ However the computation is different

$$= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - m^2$$

We can also define

* The n-th moment

$$E(X^n) = \int x^n f_X(x)\,dx$$

* The n-th central moment

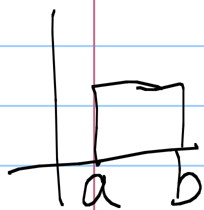
$$E((X-m)^n) = \int (x-m)^n f_X(x)\,dx$$

where m is the mean $E(X)$ #

0715

Important Conti R.V. Table 4.1 p.164

1. Uniform Random Var. with para $a < b$



$S_X = [a, b]$

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

ex: $a = 2.7$, $b = \pi$

Ex: The computer picks up a random number between $[a, b]$

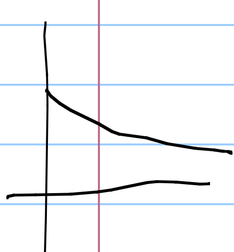
$$E(X) = \int x f_X(x)\,dx$$

$$= \int_a^b x \times \frac{1}{b-a}\,dx = \frac{a+b}{2}$$

$$Var(X) = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \times \frac{1}{b-a}\,dx$$

$$= \frac{(b-a)^2}{12} \quad \text{(exercise)}$$

2. Exponential R.V. with para $\lambda > 0$



$S = [0, \infty)$ any non-negative real number

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } 0 \le x < \infty \\ 0 & \text{otherwise} \end{cases}$$

λ can be π, 0.223, ...

Ex: Customers arrive at the average rate λ customer/per $\boxed{\text{unit time}}$. The amount of waiting time for the 1st customer is modeled by an exponential R.V.

$$E(X) = \int_0^\infty x \lambda e^{-\lambda x} dx$$

integration by part

$$= \frac{1}{\lambda} \ \boxed{\text{unit time.}}$$

Ex: average $\lambda =$ 30 customers/ $\underline{\text{hour}}$

$$E(X) = \frac{1}{\lambda} = \frac{1}{30} \ \underline{\underline{\text{hour}}}$$

unit time

We can also say
the average arrival is $\lambda = 0.5$ customers/ $\underline{\underline{\text{minute}}}$

$$E(X) = \frac{1}{\lambda} = \frac{1}{0.5} = 2 \ \underline{\underline{\text{min.}}}$$

unit time

Q: P(The first time we see a customer is > 30 min)

Ans: $P(X > 0.5)$ $\quad\|\quad$ $P(X > 30)$

$$= \int_{0.5}^\infty 30 \cdot e^{-30x} dx \quad \Big\| \quad \int_{30}^\infty 0.5 e^{-0.5x} dx$$

* Bernoulli (p) : flip a <u>single</u> coin with head prob. p.

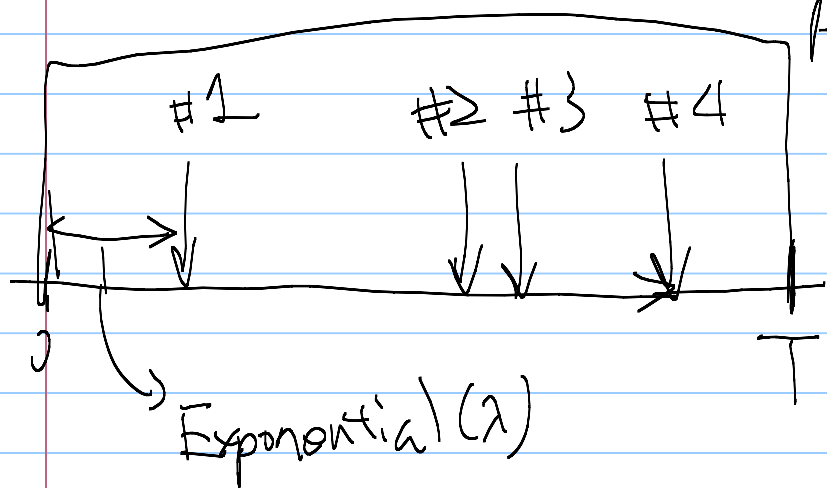
* Binomial (n, p) : flip n coins with head prob. p.
  Count the total # of heads.

* Geometric (p) : flip a coin until "head".
  Count the # of tails before the first head.

* Poisson (α) : For a given amount of time T, α is the <u>expected</u> # of arrivals in time T.
  Count the <u>actual</u> # of arrivals.

* Poisson is a limiting case of binomial.

* Exponential (λ) : λ is the <u>expected</u> # of arrivals in a <u>unit</u> time.
  Count the <u>actual</u> waiting time of the first arrival.



Poisson(α)

$\alpha = \lambda T$

E.g.

P(waiting time > 4,5) = P(No arrival in 4.5 unit time)

$$\int_{4,5}^{\infty} \lambda e^{-\lambda x}\,dx = \frac{(4,5\lambda)^0}{0!}\cdot e^{-(4,5\lambda)}$$