Discrete Random Variables

OUTLINE

- Bernoulli Distribution
- Binomial Distribution
- Geometric Distribution
- Poisson Distribution

Discrete Uniform Distribution

- Suppose *Y* can take on any integer value between *a* and *b* inclusive, each equally likely (e.g. rolling a dice, where *a*=1 and *b*=6). Then *Y* follows the discrete uniform distribution.

$$f(y) = \frac{1}{b-(a-1)} \qquad a \le y \le b$$

$$F(y) = \begin{cases} 0 & y < a \\ \dfrac{\text{int}(y)-(a-1)}{b-(a-1)} & a \le y < b \quad \text{int}(x) \equiv \text{integer portion of } x \\ 1 & y \ge b \end{cases}$$

$$E(Y) = \sum_{y=a}^{b} y\left(\frac{1}{b-(a-1)}\right) = \frac{1}{b-(a-1)}\left[\sum_{y=1}^{b} y - \sum_{y=1}^{a-1} y\right] = \frac{1}{b-(a-1)}\left[\frac{b(b+1)}{2} - \frac{(a-1)a}{2}\right] = \frac{b(b+1)-a(a-1)}{2(b-(a-1))}$$

$$E\left(Y^2\right) = \sum_{y=a}^{b} y^2\left(\frac{1}{b-(a-1)}\right) = \frac{1}{b-(a-1)}\left[\sum_{y=1}^{b} y^2 - \sum_{y=1}^{a-1} y^2\right] = \frac{1}{b-(a-1)}\left[\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)a(2a-1)}{6}\right] =$$

$$= \frac{b(b+1)(2b+1)-a(a-1)(2a-1)}{6(b-(a-1))}$$

$$\Rightarrow V(Y) = E\left(Y^2\right) - [E(Y)]^2 = \frac{b(b+1)(2b+1)-a(a-1)(2a-1)}{6(b-(a-1))} - \left[\frac{b(b+1)-a(a-1)}{2(b-(a-1))}\right]^2$$

Note : When $a = 1$ and $b = n$:

$$E(Y) = \frac{n+1}{2} \qquad V(Y) = \frac{(n+1)(n-1)}{12} \qquad \sigma = \sqrt{\frac{(n+1)(n-1)}{12}}$$

Bernoulli Distribution

- An experiment consists of one trial. It can result in one of 2 outcomes: Success or Failure (or a characteristic being Present or Absent).
- Probability of Success is p $(0<p<1)$
- Y = 1 if Success (Characteristic Present), 0 if not

$$p(y) = \begin{cases} p & y=1 \\ 1-p & y=0 \end{cases}$$

$$E(Y) = \sum_{y=0}^{1} yp(y) = 0(1-p)+1p = p$$

$$E\left(Y^2\right) = 0^2(1-p)+1^2 p = p$$

$$\Rightarrow V(Y) = E\left(Y^2\right) - \left[E(Y)\right]^2 = p - p^2 = p(1-p)$$

$$\Rightarrow \sigma = \sqrt{p(1-p)}$$

Binomial Experiment

- Experiment consists of a series of n identical trials
- Each trial can end in one of 2 outcomes: Success or Failure
- Trials are independent (outcome of one has no bearing on outcomes of others)
- Probability of Success, p, is constant for all trials
- Random Variable Y, is the number of Successes in the n trials is said to follow Binomial Distribution with parameters n and p
- Y can take on the values $y=0,1,\ldots,n$
- Notation: $Y \sim Bin(n,p)$

Binomial Distribution

Consider outcomes of an experiment with 3 Trials:

$SSS \Rightarrow y = 3 \quad P(SSS) = P(Y = 3) = p(3) = p^3$

$SSF, SFS, FSS \Rightarrow y = 2 \quad P(SSF \cup SFS \cup FSS) = P(Y = 2) = p(2) = 3p^2(1-p)$

$SFF, FSF, FFS \Rightarrow y = 1 \quad P(SFF \cup FSF \cup FFS) = P(Y = 1) = p(1) = 3p(1-p)^2$

$FFF \Rightarrow y = 0 \quad P(FFF) = P(Y = 0) = p(0) = (1-p)^3$

In General:

1) No. of ways of arranging y S^s (and $(n-y)$ F^s) in a sequence of n positions $\equiv \binom{n}{y} = \frac{n!}{y!(n-y)!}$

2) Probability of each arrangement of y S^s (and $(n-y)$ F^s) $\equiv p^y(1-p)^{n-y}$

3) $\Rightarrow P(Y = y) = p(y) = \binom{n}{y} p^y (1-p)^{n-y} \quad y = 0,1,\ldots,n$

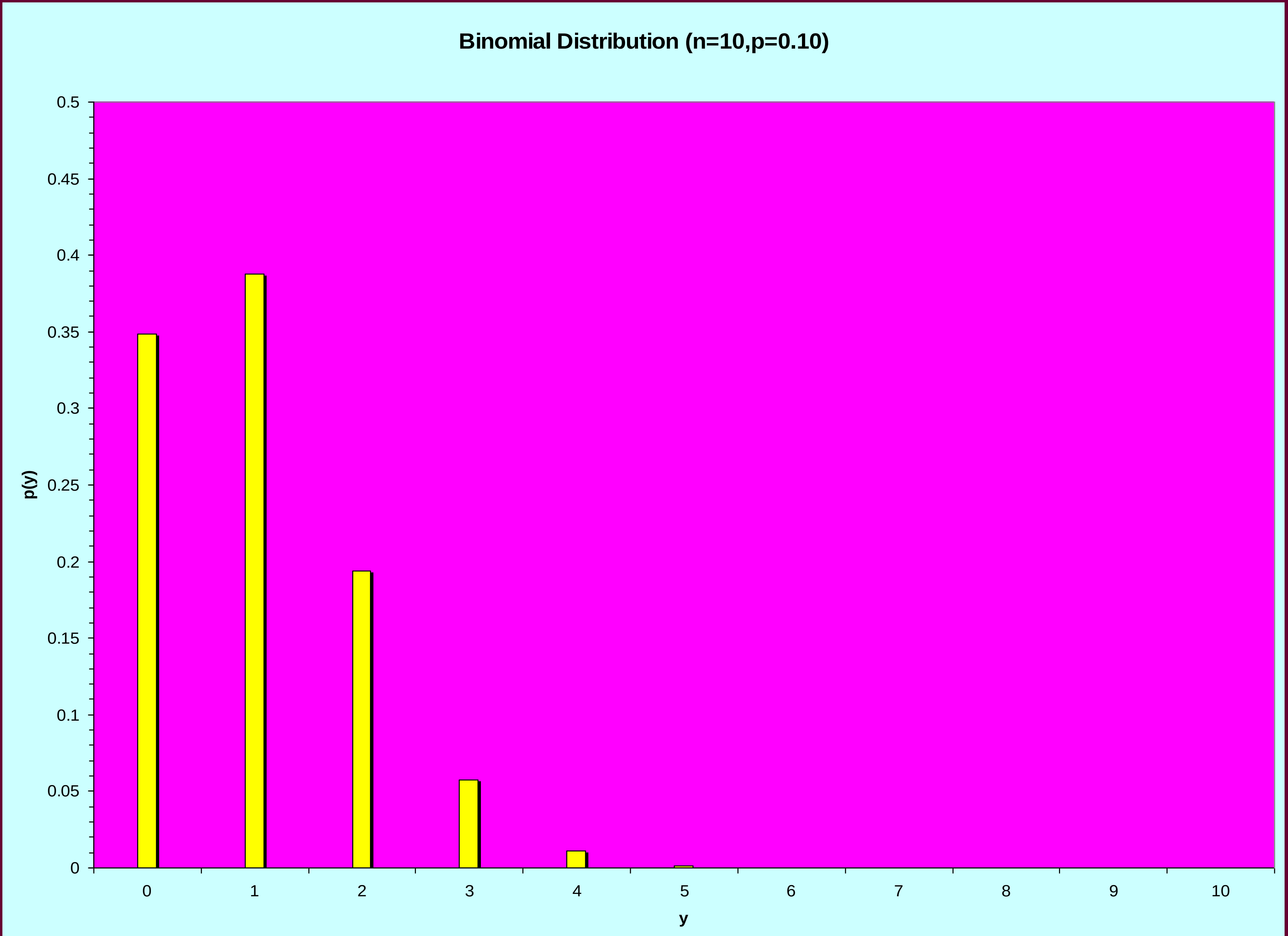
EXCEL Functions:

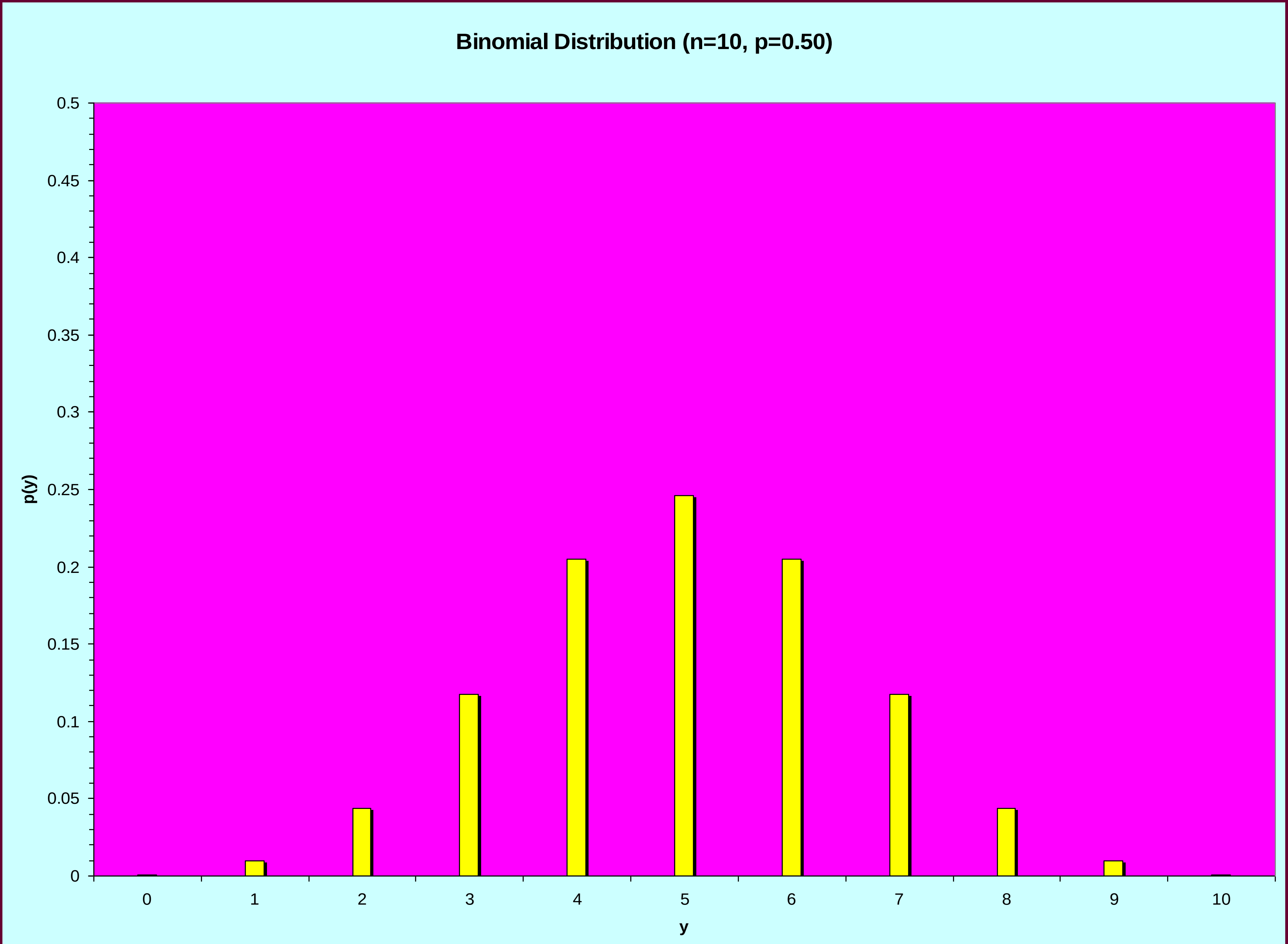
$p(y)$ is obtained by function: =BINOMDIST(y, n, p,0)

$F(y)$ is obtained by function: =BINOMDIST(y, n, p,1)

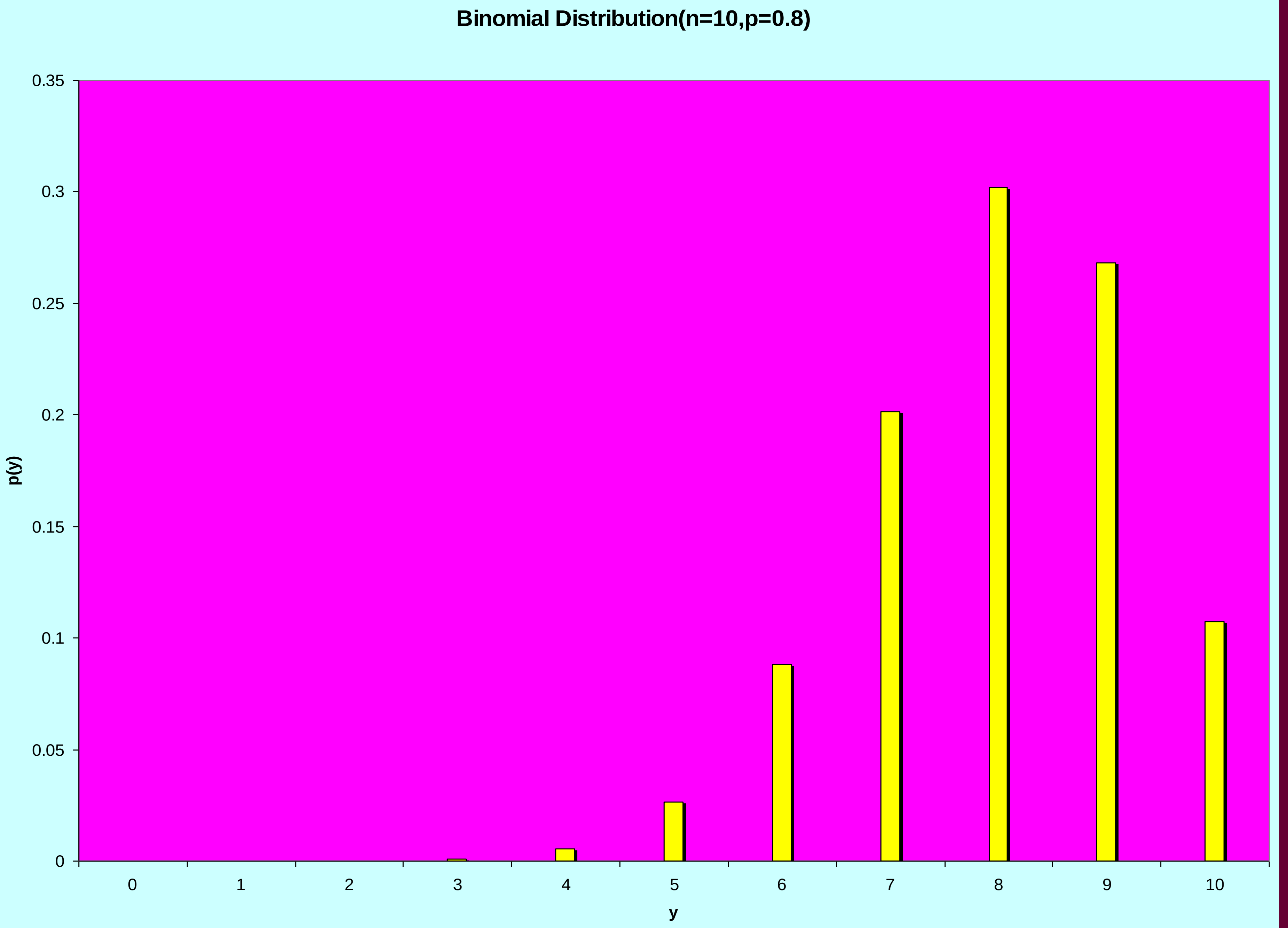
Binomial Expansion: $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$

$\Rightarrow \sum_{y=0}^{n} p(y) = \sum_{y=0}^{n} \binom{n}{y} p^y (1-p)^{n-y} = \left(p + (1-p)\right)^n = 1 \Rightarrow$ "Legitimate" Probability Distribution


Binomial Distribution (n=10,p=0.10)
p(y)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
7
8
9
10
y



Binomial Distribution (n=10, p=0.50)
p(y)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
7
8
9
10
y


Binomial Distribution(n=10,p=0.8)
p(y)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
7
8
9
10
y

Binomial Distribution – Expected Value

$$f(y) = \frac{n!}{y!(n-y)!} p^y q^{n-y} \quad y = 0,1,\ldots,n \quad q = 1-p$$

$$E(Y) = \sum_{y=0}^{n} y\left[\frac{n!}{y!(n-y)!} p^y q^{n-y}\right] = \sum_{y=1}^{n} y\left[\frac{n!}{y!(n-y)!} p^y q^{n-y}\right]$$

(Summand $= 0$ when $y = 0$)

$$\Rightarrow E(Y) = \sum_{y=1}^{n}\left[\frac{yn!}{y(y-1)!(n-y)!} p^y q^{n-y}\right] = \sum_{y=1}^{n}\left[\frac{n!}{(y-1)!(n-y)!} p^y q^{n-y}\right]$$

Let $y^* = y-1 \Rightarrow y = y^* + 1$ Note: $y = 1,\ldots,n \Rightarrow y^* = 0,\ldots,n-1$

$$\Rightarrow E(Y) = \sum_{y^*=0}^{n-1} \frac{n(n-1)!}{y^*!\left(n-(y^*+1)\right)!} p^{y^*+1} q^{n-(y^*+1)} = np\sum_{y^*=0}^{n-1} \frac{(n-1)!}{y^*!\left((n-1)-y^*\right)!} p^{y^*} q^{(n-1)-y^*} =$$

$$= np(p+q)^{n-1} = np\left(p+(1-p)\right)^{n-1} = np(1) = np$$

Binomial Distribution – Variance and S.D.

$$f(y) = \frac{n!}{y!(n-y)!} p^y q^{n-y} \quad y = 0,1,\ldots,n \quad q = 1-p$$

Note : $E(Y^2)$ is difficult (impossible?) to get, but $E(Y(Y-1)) = E(Y^2) - E(Y)$ is not :

$$E(Y(Y-1)) = \sum_{y=0}^{n} y(y-1)\left[\frac{n!}{y!(n-y)!} p^y q^{n-y}\right] = \sum_{y=2}^{n} y(y-1)\left[\frac{n!}{y!(n-y)!} p^y q^{n-y}\right]$$

(Summand $= 0$ when $y = 0,1$)

$$\Rightarrow E(Y(Y-1)) = \sum_{y=2}^{n} \frac{n!}{(y-2)!(n-y)!} p^y q^{n-y}$$

Let $y^{**} = y-2 \Rightarrow y = y^{**}+2$ Note : $y = 2,\ldots,n \Rightarrow y^{**} = 0,\ldots,n-2$

$$\Rightarrow E(Y(Y-1)) = \sum_{y^{**}=0}^{n-2} \frac{n(n-1)(n-2)!}{y^{**}!\left(n-(y^{**}+2)\right)!} p^{y^{**}+2} q^{n-(y^{**}+2)} = n(n-1)p^2 \sum_{y^{**}=0}^{n-2} \frac{(n-2)!}{y^{**}!\left((n-2)-y^{*}\right)!} p^{y^{**}} q^{(n-2)-y^{**}} =$$

$$= n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2\left(p+(1-p)\right)^{n-2} = n(n-1)p^2$$

$$\Rightarrow E(Y^2) = E(Y(Y-1)) + E(Y) = n(n-1)p^2 + np = np[(n-1)p+1] = n^2p^2 - np^2 + np = n^2p^2 + np(1-p)$$

$$\Rightarrow V(Y) = E(Y^2) - [E(Y)]^2 = n^2p^2 + np(1-p) - (np)^2 = np(1-p)$$

$$\Rightarrow \sigma = \sqrt{np(1-p)}$$

Binomial Distribution – MGF & PGF

$$M(t) = E\left(e^{tY}\right) = \sum_{y=0}^{n} e^{ty}\left[\binom{n}{y} p^{y}(1-p)^{n-y}\right] =$$

$$= \sum_{y=0}^{n}\binom{n}{y}\left(pe^{t}\right)^{y}(1-p)^{n-y} = \left(pe^{t} + (1-p)\right)^{n}$$

$$M'(t) = n\left(pe^{t} + (1-p)\right)^{n-1} pe^{t} = np\left(pe^{t} + (1-p)\right)^{n-1} e^{t}$$

$$M''(t) = np\left\{\left[(n-1)\left(pe^{t} + (1-p)\right)^{n-2} pe^{t}\right]e^{t} + \left(pe^{t} + (1-p)\right)^{n-1}\left[e^{t}\right]\right\}$$

$$\Rightarrow E(Y) = M'(0) = np\left(p(1) + (1-p)\right)^{n-1}(1) = np$$

$$\Rightarrow E\left(Y^{2}\right) = M''(0) = np\left\{\left[(n-1)\left(p(1) + (1-p)\right)^{n-2} p(1)\right](1) + \left(p(1) + (1-p)\right)^{n-1}[1]\right\} =$$

$$= np\left((n-1)p + 1\right) = n^{2}p^{2} - np^{2} + np = n^{2}p^{2} + np(1-p)$$

$$\Rightarrow V(Y) = E\left(Y^{2}\right) - \left[E(Y)\right]^{2} = n^{2}p^{2} + np(1-p) - (np)^{2} = np(1-p)$$

$$\Rightarrow \sigma = \sqrt{np(1-p)}$$

$$P(t) = E\left(t^{Y}\right) = \sum_{y=0}^{n} t^{y}\left[\binom{n}{y} p^{y}(1-p)^{n-y}\right] =$$

$$= \sum_{y=0}^{n}\binom{n}{y}(pt)^{y}(1-p)^{n-y} = \left(pt + (1-p)\right)^{n}$$

Geometric Distribution

- Used to model the number of Bernoulli trials needed until the first Success occurs (P(*S*)=*p*)
  - First Success on Trial 1 $\Rightarrow$ *S*, $y = 1 \Rightarrow p(1)=p$
  - First Success on Trial 2 $\Rightarrow$ *FS*, $y = 2 \Rightarrow p(2)=(1-p)p$
  - First Success on Trial k $\Rightarrow$ *F…FS*, $y = k \Rightarrow p(k)=(1-p)^{k-1}\, p$

$$p(y) = (1-p)^{y-1}\, p \qquad y = 1,2,\ldots$$

$$\sum_{y=1}^{\infty} p(y) = \sum_{y=1}^{\infty} (1-p)^{y-1}\, p = p\sum_{y=1}^{\infty} (1-p)^{y-1}$$

Setting $y^* = y-1$ and noting that $y = 1,2,\ldots \Rightarrow y^* = 0,1,\ldots$

$$\Rightarrow \sum_{y=1}^{\infty} p(y) = p\sum_{y^*=0}^{\infty} (1-p)^{y^*} = p\left[\frac{1}{1-(1-p)}\right] = \frac{p}{p} = 1$$

Geometric Distribution - Expectations

$$E(Y)=\sum_{y=1}^{\infty} y\left[q^{y-1}p\right]=p\sum_{y=1}^{\infty}\frac{dq^{y}}{dq}=p\frac{d}{dq}\sum_{y=1}^{\infty}q^{y}==p\frac{d}{dq}\left[q\sum_{y=1}^{\infty}q^{y-1}\right]=$$

$$=p\frac{d}{dq}\left[\frac{q}{1-q}\right]=p\left[\frac{(1-q)(1)-q(-1)}{(1-q)^{2}}\right]=\frac{p\left((1-q)+q\right)}{(1-q)^{2}}=\frac{p}{p^{2}}=\frac{1}{p}$$

$$E\left(Y(Y-1)\right)=\sum_{y=1}^{\infty} y(y-1)\left[q^{y-1}p\right]=pq\sum_{y=1}^{\infty}\frac{d^{2}q^{y}}{dq^{2}}=pq\frac{d^{2}}{dq^{2}}\sum_{y=1}^{\infty}q^{y}=pq\frac{d^{2}}{dq^{2}}\left[q\sum_{y=1}^{\infty}q^{y-1}\right]=$$

$$=pq\frac{d^{2}}{dq^{2}}\left[\frac{q}{1-q}\right]=pq\frac{d}{dq}\frac{1}{(1-q)^{2}}=pq\left(-2(1-q)^{-3}(-1)\right)=\frac{2pq}{\left(1-q\right)^{3}}=\frac{2pq}{p^{3}}=\frac{2q}{p^{2}}$$

$$\Rightarrow E\left(Y^{2}\right)=E\left(Y(Y-1)\right)+E(Y)=\frac{2q}{p^{2}}+\frac{1}{p}=\frac{2(1-p)+p}{p^{2}}=\frac{2-p}{p^{2}}$$

$$\Rightarrow V(Y)=E\left(Y^{2}\right)-\left[E(Y)\right]^{2}=\frac{2-p}{p^{2}}-\left[\frac{1}{p}\right]^{2}=\frac{2-p-1}{p^{2}}=\frac{1-p}{p^{2}}=\frac{q}{p^{2}}$$

$$\Rightarrow \sigma=\sqrt{\frac{q}{p^{2}}}$$

Geometric Distribution – MGF & PGF

$$M(t) = E\left(e^{tY}\right) = \sum_{y=1}^{\infty} e^{ty} q^{y-1} p = \frac{p}{q}\sum_{y=1}^{\infty} e^{ty} q^{y} = \frac{p}{q}\sum_{y=1}^{\infty}\left(qe^{t}\right)^{y} =$$

$$= \frac{pqe^{t}}{q}\sum_{y=1}^{\infty}\left(qe^{t}\right)^{y-1} = \frac{pe^{t}}{1-qe^{t}} = \frac{pe^{t}}{1-(1-p)e^{t}}$$

$$P(t) = E\left(t^{Y}\right) = \sum_{y=1}^{\infty} t^{y} q^{y-1} p = \frac{p}{q}\sum_{y=1}^{\infty} t^{y} q^{y} = \frac{p}{q}\sum_{y=1}^{\infty}(tq)^{y} =$$

$$= \frac{ptq}{q}\sum_{y=1}^{\infty}(tq)^{y-1} = \frac{pt}{1-tq} = \frac{pt}{1-(1-p)t}$$

Poisson Distribution

- Distribution often used to model the number of incidences of some characteristic in time or space:
  - Arrivals of customers in a queue
  - Numbers of flaws in a roll of fabric
  - Number of typos per page of text.
- Distribution obtained as follows:
  - Break down the "area" into many small "pieces" (n pieces)
  - Each "piece" can have only 0 or 1 occurrences (p=P(1))
  - Let $\lambda = np$ ≡ Average number of occurrences over "area"
  - Y ≡ # occurrences in "area" is sum of 0^s & 1^s over "pieces"
  - $Y \sim \text{Bin}(n,p)$ with $p = \lambda/n$
  - Take limit of Binomial Distribution as $n \to \infty$ with $p = \lambda/n$

Poisson Distribution - Derivation

$$p(y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} = \frac{n!}{y!(n-y)!}\left(\frac{\lambda}{n}\right)^y \left(1-\frac{\lambda}{n}\right)^{n-y}$$

Taking limit as $n \to \infty$:

$$\lim_{n\to\infty} p(y) = \lim_{n\to\infty} \frac{n!}{y!(n-y)!}\left(\frac{\lambda}{n}\right)^y \left(1-\frac{\lambda}{n}\right)^{n-y} = \frac{\lambda^y}{y!}\lim_{n\to\infty}\frac{n(n-1)\dots(n-y+1)(n-y)!}{n^y (n-y)!}\left(1-\frac{\lambda}{n}\right)^n \left(\frac{n-\lambda}{n}\right)^{-y} =$$

$$= \frac{\lambda^y}{y!}\lim_{n\to\infty}\frac{n(n-1)\dots(n-y+1)}{(n-\lambda)^y}\left(1-\frac{\lambda}{n}\right)^n = \frac{\lambda^y}{y!}\lim_{n\to\infty}\left(\frac{n}{n-\lambda}\right)\left(\frac{n-1}{n-\lambda}\right)\dots\left(\frac{n-y+1}{n-\lambda}\right)\left(1-\frac{\lambda}{n}\right)^n$$

$$\text{Note}: \lim_{n\to\infty}\left(\frac{n}{n-\lambda}\right) = \dots = \lim_{n\to\infty}\left(\frac{n-y+1}{n-\lambda}\right) = 1 \text{ for all fixed } y$$

$$\Rightarrow \lim_{n\to\infty} p(y) = \frac{\lambda^y}{y!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$

$$\text{From Calculus, we get}: \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$$

$$\Rightarrow \lim_{n\to\infty} p(y) = \frac{\lambda^y}{y!}e^{-\lambda} = \frac{e^{-\lambda}\lambda^y}{y!} \qquad y = 0,1,2,\dots$$

$$\text{Series expansion of exponential function}: e^x = \sum_{x=0}^{\infty}\frac{x^i}{i!}$$

$$\Rightarrow \sum_{y=0}^{\infty} p(y) = \sum_{y=0}^{\infty}\frac{e^{-\lambda}\lambda^y}{y!} = e^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} = e^{-\lambda}e^{\lambda} = 1 \Rightarrow \text{"Legitimate" Probability Distribution}$$

EXCEL Functions:

$p(y)$: =POISSON(y, λ,0)

$F(y)$: =POISSON(y, λ,1)

Poisson Distribution - Expectations

$$f(y) = \frac{e^{-\lambda}\lambda^y}{y!} \qquad y = 0,1,2,\ldots$$

$$E(Y) = \sum_{y=0}^{\infty} y\left[\frac{e^{-\lambda}\lambda^y}{y!}\right] = \sum_{y=1}^{\infty} y\left[\frac{e^{-\lambda}\lambda^y}{y!}\right] = \sum_{y=1}^{\infty} \frac{e^{-\lambda}\lambda^y}{(y-1)!} = \lambda e^{-\lambda}\sum_{y=1}^{\infty}\frac{\lambda^{y-1}}{(y-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

$$E\left(Y(Y-1)\right) = \sum_{y=0}^{\infty} y(y-1)\left[\frac{e^{-\lambda}\lambda^y}{y!}\right] = \sum_{y=2}^{\infty} y(y-1)\left[\frac{e^{-\lambda}\lambda^y}{y!}\right] = \sum_{y=2}^{\infty}\frac{e^{-\lambda}\lambda^y}{(y-2)!} =$$

$$= \lambda^2 e^{-\lambda}\sum_{y=2}^{\infty}\frac{\lambda^{y-2}}{(y-2)!} = \lambda^2 e^{-\lambda}e^{\lambda} = \lambda^2$$

$$\Rightarrow E\left(Y^2\right) = E\left(Y(Y-1)\right) + E(Y) = \lambda^2 + \lambda$$

$$\Rightarrow V(Y) = E\left(Y^2\right) - [E(Y)]^2 = \lambda^2 + \lambda - [\lambda]^2 = \lambda$$

$$\Rightarrow \sigma = \sqrt{\lambda}$$

Poisson Distribution – MGF & PGF

$$M(t) = E\left(e^{tY}\right) = \sum_{y=0}^{\infty} e^{ty}\left[\frac{e^{-\lambda}\lambda^{y}}{y!}\right] = \sum_{y=0}^{\infty} \frac{e^{-\lambda}\left(\lambda e^{t}\right)^{y}}{y!} =$$

$$= e^{-\lambda}\sum_{y=0}^{\infty} \frac{\left(\lambda e^{t}\right)^{y}}{y!} = e^{-\lambda}e^{\lambda e^{t}} = e^{\lambda\left(e^{t}-1\right)}$$

$$P(t) = E\left(t^{Y}\right) = \sum_{y=0}^{\infty} t^{y}\left[\frac{e^{-\lambda}\lambda^{y}}{y!}\right] = \sum_{y=0}^{\infty} \frac{e^{-\lambda}(\lambda t)^{y}}{y!} =$$

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

THANK YOU

Submitted by:
Rakesh Kumar,
(Deptt. Of Mathematics),
P.G.G.C.G, Sec-11,
Chandigarh.