

Stat03-Poisson.wmxm: Poisson(λ) Distribution

TABLE OF CONTENTS

Preface	1
References	1
Discrete Distributions Defined	3
What Is a Discrete Random Variable?	3
Mean and Variance of Discrete Data Set	5
The Discrete Poisson(λ) Distribution	5
pdf_poisson (n, λ), mean_poisson (λ)	7
std_poisson(λ), cdf_poisson (n, λ)	10
cdf_poisson ($k1, \lambda$) Use Example 1	11
cdf_poisson Use Example 2, Approximating the Binomial Distribution	13
quantile_poisson (q, λ)	14
poissonCalc ($k1, \lambda$)	15
FS Ex 1: Number of Cars Through Intersection 12 noon - 1 pm	17
FS Ex 2: Number of Particles Emitted by Source over del_t	19
Statology 1: Calls per Hour at a Call Center	21
Statology 2: Number of Restaurant Customers per Day	22
Statology 3: Number of Website Visitors per Hour	22
Statology 4: Number of Bankruptcies Filed per Month	23
Statology 5: Number of Network Failures per Week	24
Number of Beam Particles per Pulse	24
[RS] Prob. 3.30	25
Barlow Problem 3.3	25
Barlow Problem 3.4	26
Barlow Problem 3.5	26
random_poisson (λ), random_poisson (λ, n)	27
Random Sample Size $m = 10$ Simulations, $\lambda = 15$	27
Random Sample Size $m = 100$ Simulations, $\lambda = 15$	27
discrete_freq (data)	28
Random Sample Size $m = 1000$ Simulations, $\lambda = 15$	30

1 Preface

In Stat03-Poisson.wmxm we discuss the discrete Poisson(λ) probability distribution and its application, using Maxima tools and methods. This is the third worksheet in my Statistics with Maxima section.

Edwin L. (Ted) Woollett
<https://home.csulb.edu/~woollett/>
 April 18, 2024

2 References

In our series Statistics with Maxima we have used some examples and explanations (with much editing and additions) from:

Roger Barlow, Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences, John Wiley, 1989,

Ch. 3 Reagle & Salvatore [RS], Statistics and Econometrics, 2nd ed, Schaum's Outlines, 2011, McGraw Hill,

Ch. 8 Fred Senese [FS], Symbolic Mathematics for Chemists: A Guide for Chemists, 2019, Wiley,

Louis Lyons, Statistics for Nuclear and Particle Physics, 1986, Cambridge Univ. Press,

Luca Lista, 'Statistical Methods for Data Analysis in Particle Physics', Lecture Notes in Physics 909, 2016, Springer-Verlag,

Frederick James, 'Statistical Methods in Experimental Physics', 2nd ed., 2006, World Scientific.

In this Stat03-Poisson.wmx worksheet we use examples from:

<https://www.statology.org/poisson-distribution-real-life-examples/>

<https://www.statology.org/poisson-distribution-calculator/>

<https://online.stat.psu.edu/stat414/lesson/12>

```
(%i5) load (descriptive);
load (distrib);
fpprintprec : 6$
ratprint : false$
logexpand : all$
```

```
(%o1)
```

```
C:/maxima-5.43.2/share/maxima/5.43.2/share/descriptive/descriptive.mac
```

```
(%o2)
```

```
C:/maxima-5.43.2/share/maxima/5.43.2/share/distrib/distrib.mac
```

Homemade functions fill, head, tail, Lsum are useful for looking at long lists.

```
(%i13) fill ( aL) := [ first (aL), last (aL), length (aL) ]$
declare (fill, evfun)$
head(L) := if listp (L) then rest (L, - (length (L) - 3) ) else
error("Input to 'head' must be a list of expressions ")$
declare(head, evfun)$
tail (L) := if listp (L) then rest (L, length (L) - 3 ) else
error("Input to 'tail' must be a list of expressions ")$
declare(tail, evfun)$
Lsum (aList) := apply ("+", aList)$
declare (Lsum, evfun)$
```

3 *Discrete Distributions Defined*

From www.investopedia.com:

"A discrete probability distribution counts occurrences that have countable or finite outcomes.

Discrete distributions contrast with continuous distributions, where outcomes can fall anywhere on a continuum.

Common examples of discrete distribution include the binomial, Poisson, and Bernoulli distributions.

These distributions often involve statistical analyses of "counts" or "how many times" an event occurs.

In finance, discrete distributions are used in options pricing and forecasting market shocks or recessions."

From <http://www.stat.yale.edu/Courses/1997-98/101/ranvar.htm>:

"If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten."

4 *What Is a Discrete Random Variable?*

Quoting Luca Lista (Sec. 1.1):

"Many processes in nature have uncertain outcomes. This means that their result cannot be predicted before the process occurs. A random process is a process that can be reproduced, to some extent, within some given boundary and initial conditions, but whose outcome is uncertain. This situation may be due to insufficient information about the process intrinsic dynamics which prevents to predict its outcome, or lack of sufficient accuracy in reproducing the initial conditions in order to ensure its exact reproducibility. Some processes like quantum mechanics phenomena have intrinsic randomness. This will lead to possibly different outcomes if the experiment is repeated several times, even if each time the initial conditions are exactly reproduced, within the possibility of control of the experimenter. Probability is a measurement of how favored one of the possible outcomes of such a random process is compared with any of the other possible outcomes."

A coin toss is "random" because we are ignorant of the 'initial conditions'. Repeated trials tell us something about how those initial conditions vary between trials

For the purposes of calculating things for experimental physics, we need physical probability. In particular we need 'frequentist probability':

"Probability is the frequency with which a particular outcome occurs in repeated trials."

$$P = (\text{number of occasions on which that outcome occurs}) / (\text{total number of measurements}).$$

Quoting L. Lyons, Sec. 2.1:

"In many situations we deal with experiments in which the essential circumstances are kept constant, and yet repetitions of the experiment produce different results. Thus the result of an individual measurement or trial may be unpredictable, and yet the possible results of a series of such measurements have a well defined distribution."

What about events that can't be repeated? They don't have probabilities.

Quoting [RS] Sec. 3.3:

"A random variable is a variable whose values are associated with some probability of being observed. A discrete (as opposed to continuous) random variable is one that can assume only finite and distinct values. The set of all possible values of a random variable and its associated probabilities is called a probability distribution. The sum of all probabilities equals 1."

Quoting

https://saylordotorg.github.io/text_introductory-statistics/s08-discrete-random-variables.html,

"The probability distribution of a discrete random variable X is a listing of each possible value x taken by X along with the probability $P(x)$ that X takes that value in one trial of the experiment.

The mean μ of a discrete random variable X is a number that indicates the average value of X over numerous trials of the experiment. It is computed using the formula $\mu = \sum x P(x)$.

The variance σ^2 and standard deviation σ of a discrete random variable X are numbers that indicate the variability of X over numerous trials of the experiment. They may be computed using the formula $\sigma^2 = (\sum x^2 P(x)) - \mu^2$, taking the square root to obtain σ ."

5 Mean and Variance of Discrete Data Set

Consider a data set that contains M unique discrete values x_k , and assume the value x_k occurs with frequency f_k . Let N equal the sum of the frequencies.

$$N = \sum (f_k, k, 1, M).$$

The mean

$$\langle x \rangle = \sum (f_k x_k, k, 1, M) / N.$$

The variance

$$\text{Var}(x) = \sum (f_k (x_k - \langle x \rangle)^2, k, 1, M) / N.$$

The standard deviation is the square root of the variance.

6 The Discrete Poisson(λ) Distribution

Recall our definition of the Binomial (n, p) distribution in Stat02-Binomial.wmx:

The binomial distribution is used to find the probability of k 'successes' $P_b(k; n, p)$ in n trials of the same experiment when (1) there are only two possible and mutually exclusive outcomes, (2) the n trials are independent, and (3) the probability of 'success', p , remains constant in each trial.

Quoting Roger Barlow, Sec. 3.3,

"The binomial distribution describes cases where particular outcomes occur in a certain number of trials, n [of the same experiment/measurement]."

"The Poisson distribution describes cases where there are still particular outcomes [observations] but no idea of the number of trials [or measurements carried out under the same conditions]; instead these are sharp events occurring in a continuum. For example, during a thunderstorm there will be a definite number of flashes of lightning (sharp events), but it is meaningless to ask how often there was **not** a flash."

"A Geiger counter placed near a radioactive source will produce definite clicks, but not definite non-clicks. If in such an experiment one knows that the average number of events is, say, ten a minute, then in a particular minute one expects on average ten events, though intuitively one feels that nine or eleven would be unremarkable... but suppose there were five or fifteen? Is that compatible, or has something changed? We need to know the probability of obtaining a particular number of events, given the average number. This can be analysed by taking the limit of the binomial distribution, in which the number of tries [of the same measurement/trial], n , becomes large while at the same time the probability p becomes small, with their product constant."

"Suppose that on average λ events would be expected to occur in some interval. Split the interval up into n very small equal sections, so small that the chance of getting two events in one section can be discounted. Then the probability that a given section contains an event is $p = \lambda/n$. The probability that there will be k events in the n sections of the interval is given by the binomial formula:

$$P_b(k; n, p) = \left(\frac{n!}{k! \cdot (n - k)!} \right) \cdot p^k \cdot (1 - p)^{(n - k)}$$

$$= \left(\frac{n!}{k! \cdot (n - k)!} \right) \cdot (\lambda/n)^k \cdot (1 - \lambda/n)^{(n - k).}"$$

As $n \rightarrow \infty$ with k finite, this expression becomes the Poisson probability formula, the probability of obtaining exactly k events if the mean expected number is λ : [where $\lambda = p \cdot n$]

$$P(k, \lambda) = \exp(-\lambda) \cdot \lambda^k / k! .$$

For a derivation in detail see (where $k \rightarrow x$): <https://online.stat.psu.edu/stat414/lesson/12/12.1>

Quoting <https://online.stat.psu.edu/stat414/lesson/12/12.1>, with some light editing,

"Let the discrete random variable k denote the number of times an event occurs in an interval of time (or space). Then k may be a Poisson random variable with $k = 0, 1, 2, \dots$ "

"Examples:

- 1.) Let k equal the number of typos on a printed page. (This is an example of an interval of space — the space being the printed page.)
- 2.) Let k equal the number of cars passing through the intersection of Allen Street and College Avenue in one minute. (This is an example of an interval of time — the time being one minute.)
- 3.) Let k equal the number of Alaskan salmon caught in a squid driftnet. (This is again an example of an interval of space — the space being the squid driftnet.)
- 4.) Let k equal the number of customers at an ATM in 10-minute intervals.
- 5.) Let k equal the number of students arriving during office hours."

"Poisson Random Variable

If k is a Poisson random variable, then the probability mass function [probability distribution function] is:

$$f(k) = \exp(-\lambda) * \lambda^k / k!$$

for $k = 0, 1, 2, \dots$ and $\lambda > 0$, where λ will be shown later to be both the mean and the variance of the distribution. Also, note that there are (theoretically) an infinite number of possible Poisson distributions. Any specific Poisson distribution depends on the parameter λ ."

6.1 pdf_poisson (k, λ), mean_poisson (λ)

pdf_poisson (k, λ) is the probability of counting exactly k events when the average number of events observed in a trial is λ and when the distribution of events is described by the Poisson distribution. In particular, the outcome of each trial is completely independent of the result observed in the previous trial.

mean_poisson (λ) returns λ .

```
(%i15) λ;  
      mean_poisson (λ);
```

```
(%o14) λ
```

```
(%o15) λ
```

Quoting RS in Sec. 3.4: (They shift language from 'event' to 'success'.)

"The Poisson distribution is another discrete probability distribution. It is used to determine the probability of a designated number of 'successes' per unit of time, when the events or successes are independent, and the average number of successes per unit of time remains constant. Then

$$P(k) = \lambda^k \exp(-\lambda) / k!$$

where k is the designated number of successes, $P(k)$ is the probability of k successes ($k = 0, 1, 2, 3, \dots$), and λ is the average number of successes per unit of time."

The mean is λ , the variance is λ , and one standard deviation is $\sqrt{\lambda}$.

Quoting Fred Senese in Sec. 8.2.2:

"When the sample size is large, the discrete probability density often converges to a smoothly varying 'limiting distribution'. Limiting distributions can be used to compute probabilities directly."

"Suppose we are interested in measuring a 'count': [for example] the number of photons emitted by a chemiluminescent reaction, or [as another example] the number of particles emitted by a radioactive substance." Each photon or particle which appears is the result of a random event, with a random interval of time between successive events, "but they have a fixed [characteristic] average 'rate' r " over time; r has units (number/time)."

"If we measure the number of [such] events [which] occur over [a specified] time t [getting a 'count' each time we repeat the measurement], the 'average count' λ will be:

$$\lambda = r * t = (\text{number/time}) * \text{time} = \text{number} = \text{dimensionless},$$

after making a large number n of measurements of this average count."

"Inverting this equation, we have an experimental value for the 'fixed average rate' of arrival $r = \lambda / t$."

"[It can be shown that] the probability of actually finding exactly k events in any given measurement is given by $P(k, \lambda)$:

$$P(k, \lambda) = \lambda^k * \exp(-\lambda) / k!$$

known as the 'Poisson distribution', an example of a 'probability distribution'."

"Since k is an integer (the number of counts observed in a trial), $P(k, \lambda)$ is a discrete probability distribution, and it can be proved that for a data series x_L of events described by the Poisson distribution, $\langle k \rangle = \lambda$, $\text{var}[k] = \lambda$, and hence $\sigma[k] = \sqrt{\lambda} = \lambda^{1/2}$."

Quoting Luca Lista in 'Statistical Methods for Data Analysis in Particle Physics',
Lecture Notes in Physics 909:

"A discrete variable k is said to be a Poissonian variable if it obeys the Poisson distribution defined below for a given value of the parameter λ :

$$P(k; \lambda) = \lambda^k * \exp(-\lambda)/k! "$$

```
(%i16) mean_poisson (15);
```

```
(%o16) 15
```

"For example, suppose we want to know how likely it will be for $k = 10$ molecules to arrive over a $1\mu\text{s}$ interval [of time] at the surface of a sensor." If the average rate r has been measured to be $r = 15$ molecules/ μs , and we let $t = 1\mu\text{s}$, then $\lambda = r * t = (15 \text{ molecules}/\mu\text{s}) * 1 \mu\text{s} = 15 \text{ molecules} = \text{average count over } 1 \mu\text{s}$. We then use the poisson probability distribution to determine the probability $P(10)$ of actually finding $k = 10$ molecules arriving in $1 \mu\text{s}$ in a given measurement:

```
(%i18) ( (15^10) / factorial (10) ) * exp (- 15), numer;  
pdf_poisson (10, 15), numer;
```

```
(%o17) 0.0486108
```

```
(%o18) 0.0486108
```

The chance of measuring exactly 10 molecules arriving in a given micro-second is about 5%.

"Let's plot the Poisson distribution in this case $\lambda = 15$. After building a list of 30 [$k, P(k, \lambda)$] 'points' with varying n and with $\lambda = 15$, we plot them using the impulses style: "

The maximum value will be for $n = 15$.

```
(%i19) pdf_poisson (15, 15), numer;
```

```
(%o19) 0.102436
```

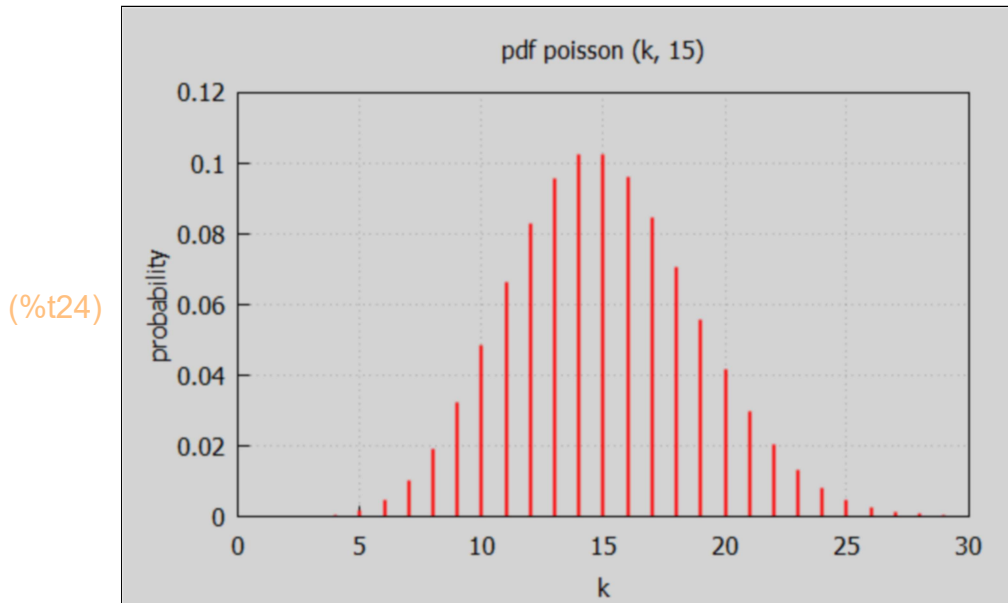
```
(%i23) mypoints : makelist ( [k, float (pdf_poisson (k, 15))] , k, 1, 30)$  
fill (mypoints);  
head (mypoints);  
tail (mypoints);
```

```
(%o21) [[1,4.58853 10-6],[30,2.21137 10-4],30]
```

```
(%o22) [[1,4.58853 10-6],[2,3.4414 10-5],[3,1.7207 10-4]]
```

```
(%o23) [[28,8.55061 10-4],[29,4.42273 10-4],[30,2.21137 10-4]]
```

```
(%i24) wxdraw2d ( xrange = [0, 30], yrange = [0, 0.12],
  points_joined = impulses,
  title = "pdf poisson (k, 15)",
  xlabel = "k", ylabel = "probability", grid = true,
  background_color = light_gray,
  line_width = 2, color = red,
  points (mypoints) )$
```



"The Poisson distribution is discrete (the count k is an integer). It isn't quite symmetric; it tails off faster on the left than it does on the right. The mean of the distribution is λ . We can describe the 'width' of the distribution in terms of the root-mean-square [rms] deviation from the mean (which is called the 'standard deviation' [sd] when there are a large number of trials)."

Since k is an integer (the number of counts observed in a trial), $P(k, \lambda)$ is a discrete probability distribution, and it can be proved that for a data series x_L of events described by the Poisson distribution, $\langle k \rangle = \lambda$, $\text{var}[k] = \lambda$, and hence $\sigma[k] = \sqrt{\lambda} = \lambda^{1/2}$.

" The Poisson distribution $P(k, \lambda)$ is a built-in Maxima function $\text{pdf_poisson}(k, \lambda) = \text{probability of counting exactly } k \text{ events}$ ". The pdf prefix stands for 'probability distribution function'.

6.2 $\text{std_poisson}(\lambda)$, $\text{cdf_poisson}(k1, \lambda)$

Quoting [RS]:

"The width of the distribution is characterized by the standard deviation σ . The interval $[\lambda - \sigma, \lambda + \sigma]$ encompasses roughly two-thirds of all of the data. For the Poisson distribution, $\sigma = \sqrt{\lambda} = \text{'standard deviation of the Poisson distribution'}$ " and we can use the Maxima function $\text{std_poisson}(\lambda)$.

```
(%i26) std_poisson (15);
      sigma : std_poisson (15), numer;
(%o25)  $\sqrt{15}$ 
(sigma) 3.87298
```

"The 'cumulative distribution function' `cdf_poisson (k1, λ)` returns the sum of all probabilities for 0 to k1 counts, which computes the probability $P(k \leq k1)$ that the actual count observed will lie in the interval $[0, k1]$. Let's use it to compute the probability of getting a count [somewhere in the interval] between $\lambda - \sigma$ and $\lambda + \sigma$."

The number we want is the 'area under the curve' from $\mu - \sigma$ to $\mu + \sigma$, which is given by the difference:

```
(%i27) cdf_poisson (15 + sigma, 15) - cdf_poisson (15 - sigma, 15), numer;
(%o27) 0.63472
```

which is 'roughly two thirds'.

Let's compare `cdf_poisson (k1, λ)` with `sum (pdf_poisson (k, λ), k, 0, k1)`.

```
(%i28) compare (kk,  $\lambda$ ) :=
      [kk, float (cdf_poisson (kk,  $\lambda$ )), float (sum (pdf_poisson (k,  $\lambda$ ), k, 0, kk))];
(%i29) compare (5, 15);
(%o29) [5,0.00279243,0.00279243]

(%i30) for k : 12 thru 18 do print (compare (k, 15))$
      [12,0.267611,0.267611]
      [13,0.363218,0.363218]
      [14,0.465654,0.465654]
      [15,0.56809,0.56809]
      [16,0.664123,0.664123]
      [17,0.748859,0.748859]
      [18,0.819472,0.819472]
```

6.2.1 `cdf_poisson (k1, λ)` Use Example 1

Quoting Example 12.2 on: <https://online.stat.psu.edu/stat414/lesson/12/12.2>
with some light editing:

"Let k equal the number of typos on a printed page with a mean of 3 typos per page."

Recall that the mean (ie., the average) number of counts is the same as the parameter λ , so in this example $\lambda = 3$.

"A.) What is the probability that a randomly selected page has at least one typo on it?
We can find the requested probability directly from the pdf."

The probability that k is at least equal to 1 is:

$P(k \geq 1) =$ probability the number of typos actually found on a random page is $k = 1, 2, 3, \dots$
but $P(k < 1) + P(k \geq 1) = 1$, so

$P(k \geq 1) = 1 - P(k < 1) = 1 - \text{cdf_poisson}(0, \lambda)$, but
 $\text{cdf_poisson}(0, \lambda) = \text{pdf_poisson}(0, \lambda)$.

```
(%i32) cdf_poisson(0, 15), numer;  
pdf_poisson(0, 15), numer;
```

```
(%o31) 3.05902 10-7
```

```
(%o32) 3.05902 10-7
```

Hence $P(k \geq 1) = 1 - \text{pdf_poisson}(0, \lambda)$.

```
(%i33) 1 - pdf_poisson(0, 3), numer;
```

```
(%o33) 0.950213
```

"That is, there is just over a 95% chance of finding at least one typo on a randomly selected page when the average number of typos per page is 3."

"B.) What is the probability that a randomly selected page has at most one typo on it?"

$P(k \leq 1) = P(k = 0) + P(k = 1) = \text{cdf_poisson}(1, \lambda)$, so

```
(%i34) cdf_poisson(1, 3), numer;
```

```
(%o34) 0.199148
```

"That is, there is just under a 20% chance of finding at most one typo on a randomly selected page when the average number of typos per page is 3."

"C.) What is the probability that three randomly selected pages have more than eight typos on it?"

"Solving this problem involves taking one additional step. Recall that n denotes the number of typos on one printed page."

"Let's define a new random variable y that equals the number of typos on three printed pages. If the mean of k is 3 typos per page, then the mean of y is:

$$\lambda y = 3 \text{ typos per one page} * 3 \text{ pages} = 9 \text{ typos per 3 pages.}''$$

"Finding the desired probability then involves finding

$$P(y > 8),'' \text{ and}$$

$$P(y \leq 8) + P(y > 8) = 1, \text{ so}$$

$$P(y > 8) = 1 - P(y \leq 8) = 1 - \text{cdf_poisson}(8, \lambda y) = 1 - \text{cdf_poisson}(8, 9).$$

```
(%i35) 1 - cdf_poisson(8, 9), numer;
```

```
(%o35) 0.544347
```

"..... there is a 54.4% chance that three randomly selected pages would have more than eight typos on it."

6.2.2 cdf_poisson(k1, λ) Use Example 2, Approximating the Binomial Distribution

Quoting loosely from Example 12.3 in

<https://online.stat.psu.edu/stat414/lesson/12/12.4>

"Five percent (5%) of Christmas tree light bulbs manufactured by a company are defective. The company's Quality Control Manager is quite concerned and therefore randomly samples 100 bulbs coming off of the assembly line. "

"Let k denote the number in the sample that are defective. What is the probability that the sample contains at most three defective bulbs?"

Every sample of $n = 100$ light bulbs from the assembly line has the same average number of defective bulbs (we assume). The average number of defective light bulbs in a sample of 100 bulbs is $\lambda = n * p = 100 * 0.05 = 5$. Let k = the number of defective bulbs found in a sample of 100 bulbs for which $\lambda = 5$.

We are asked to find $P(k \leq 3) = P(0) + P(1) + P(2) + P(3) = \text{cdf_poisson}(3, 5)$

```
(%i36) cdf_poisson(3, 5), numer;
```

```
(%o36) 0.265026
```

"... there is a 26.5% chance that a randomly selected batch of 100 bulbs will contain at most 3 defective bulbs."

Using the binomial distribution $P(k \leq 3) = \text{cdf_binomial}(3, 100, 0.05)$."

```
(%i37) cdf_binomial (3, 100, 0.05);
(%o37) 0.257839
```

which is quite close to 0.265 gotten using the Poisson distribution with $\lambda = n \cdot p = 100 \cdot 0.05 = 5$.

"But, if you recall the way that we derived the Poisson distribution,... we started with the binomial distribution and took the limit as n approached infinity. So, it seems reasonable then that the Poisson distribution would serve as a reasonable approximation to the binomial distribution when your value of n trials is large (and therefore, p is small). "

You then approximate Binomial (n, p) with Poisson ($\lambda = n \cdot p$).

"It is important to keep in mind that the Poisson approximation to the binomial distribution works well only when n is large and p is small. In general, the approximation works well if $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $p \leq 0.1$."

Loosely quoting [RS], p. 40:

"The Poisson distribution can be used as an approximation to the binomial distribution when n is large (a large number of independent trials or a large sample) and p or $(1-p)$ is small: say, $n \geq 30$ and $n \cdot p < 5$ or $n \cdot (1-p) < 5$."

6.3 quantile_poisson (q, λ)

`quantile_poisson(q, λ)` returns the number of events counted with probability q , as long as q is an element of the numbers produced by `cdf_poisson(k, λ)`.

From the Maxima manual:

`quantile_poisson(q, λ)` returns the q -quantile of a $\text{Poisson}(\lambda)$ random variable, with $\lambda > 0$; in other words, this is the inverse of `cdf_poisson`. Argument q must be an element of $[0, 1]$.

```
(%i38) cdf_poisson (13, 15), numer;
(%o38) 0.363218
```

```
(%i39) quantile_poisson (% , 15);
(%o39) 13
```

```
(%i40) for k:0 thru 15 do (
  qq : float (cdf_poisson (k, 15) ),
  print (k, qq, quantile_poisson (qq, 15) )) $
0 3.05902 10-7 0
1 4.89444 10-6 1
2 3.93084 10-5 2
3 2.11379 10-4 3
4 8.56641 10-4 4
5 0.00279243 5
6 0.0076319 6
7 0.0180022 7
8 0.0374465 8
9 0.0698537 9
10 0.118464 10
11 0.184752 11
12 0.267611 12
13 0.363218 13
14 0.465654 14
15 0.56809 15
```

6.4 poissonCalc (k1, λ)

Let's first use the online poisson distribution probability calculator at the webpage:
<https://www.statology.org/poisson-distribution-calculator/>

Here is the output for the case $n = 5$, $\lambda = 15$.

$P(X = 5)$: 0.00194

$P(X < 5)$: 0.00086

$P(X \leq 5)$: 0.00279

$P(X > 5)$: 0.99721

$P(X \geq 5)$: 0.99914

We are going to check these results using the Maxima functions, and in the process get the path needed to define our own Maxima function, poissonCalc (k, λ).

```
(%i41) fpprintprec : 5$
```

```
(%i42) pdf_poisson (5, 15), numer;
(%o42) 0.0019358

(%i43) cdf_poisson (4, 15), numer;
(%o43) 8.5664 10-4

(%i44) cdf_poisson (5, 15), numer;
(%o44) 0.0027924

(%i45) 1 - cdf_poisson (5, 15), numer;
(%o45) 0.99721

(%i46) 1 - cdf_poisson (4, 15), numer;
(%o46) 0.99914
```

These answers confirm those supplied by the online calculator. We will use the same commands in our Maxima function poissonCalc (n, λ).

```
(%i47) poissonCalc (kk, λ) :=
block (
  print (" "),
  print (sconcat (" For k = ", kk, ", λ = ", λ, ", the probabilities are:")),
  print (" "),
  print (sconcat (" P ( k = ", kk, " ) : ", float (pdf_poisson (kk, λ)) )),
  print (sconcat (" P ( k < ", kk, " ) : ", float (cdf_poisson (kk - 1, λ)) )),
  print (sconcat (" P ( k <= ", kk, " ) : ", float (cdf_poisson (kk, λ)) )),
  print (sconcat (" P ( k > ", kk, " ) : ", float (1 - cdf_poisson (kk, λ)) )),
  print (sconcat (" P ( k >= ", kk, " ) : ", float (1 - cdf_poisson (kk - 1, λ)) )),
  done )$
```

Here we try out this function:

```
(%i48) poissonCalc (5, 15)$
```

For k = 5, λ = 15, the probabilities are:

```
P ( k = 5 ) : 0.0019358
P ( k < 5 ) : 8.5664e-4
P ( k <= 5 ) : 0.0027924
P ( k > 5 ) : 0.99721
P ( k >= 5 ) : 0.99914
```

So our homemade function returns what we expect in this example.

6.5 FS Ex 1: Number of Cars Through Intersection 12 noon - 1 pm

This is a combination of 'Problem 2' and 'Problem 3' on Fred Senese's Ch. 8 discrete pdf worksheet. Between noon and one o'clock, an average of 25 cars pass through a particular intersection.

A) What is the chance that exactly 25 cars pass through the intersection on a particular day?

In this example each trial is an observation of the number of cars which pass through the chosen intersection (in the hour 12 noon to 1 pm), each day's trial is completely independent of any other day's trial (we assume) and the average number has been measured to be $\lambda = 25$ cars. We have then an experimental value for the 'fixed average rate' of passage of cars through the intersection given by $r = \lambda/\Delta t = 25$ cars/hour, which is assumed to be the same each day.

`pdf_poisson(k, λ)` is the probability of counting exactly k events when the average number of events observed in a trial is λ and when the distribution of events is described by the Poisson distribution. In particular, the outcome of each trial is completely independent of the result observed in the previous trial.

```
(%i49) pdf_poisson(25, 25), numer;
```

```
(%o49) 0.079523
```

So about an 8% chance exactly 25 cars pass through the specified intersection on any particular day.

B) What is the chance that fewer than 25 cars pass through the intersection?

C) What is the chance that more than 25 cars pass through the intersection?

Both parts B and C can be answered with our function `poissonCalc(k1, λ)`.

```
(%i50) poissonCalc(25, 25)$
```

For $k = 25$, $\lambda = 25$, the probabilities are:

$P(k = 25) : 0.079523$

$P(k < 25) : 0.4734$

$P(k \leq 25) : 0.55292$

$P(k > 25) : 0.44708$

$P(k \geq 25) : 0.5266$

In words, the chance that fewer than 25 cars pass through the intersection on a given day is 47% and the chance that more than 25 cars pass through the intersection on a given day is 45%.

D) There is a 95% chance that anywhere from zero to how many cars will pass through the intersection?

```
(%i51) quantile_poisson (0.95, 25);
(%o51) 33
```

Bearing in mind that `quantile_poisson (q, λ)` is only the inverse of `cdf_poisson (k, λ)` if q belongs to the discrete set of numbers produced by `cdf_poisson (k, λ)` for $k = 0, 1, 2, \dots$, we need to explore the values of `cdf_poisson (n, 25)` for k near the integer 33.

```
(%i54) cdf_poisson (32, 25), numer;
      cdf_poisson (33, 25), numer;
      cdf_poisson (34, 25), numer;
(%o52) 0.92854
(%o53) 0.95022
(%o54) 0.96616
```

So we can be 95% confident that anywhere from zero to 33 cars will pass through the chosen intersection during the noon hour on any given day (the actual count observed will lie in the interval $[0, 33]$).

E) Make an impulse plot of the probability density distribution for the cars passing through the intersection in the previous problem. Show the probabilities for 0 to 50 cars in your plot.

The average is 25 and the standard deviation is the square root of the average (for a Poissonian variable). We will add the plot of `pdf_normal (x, 25, 5)` to the plot.

```
(%i56) mean_poisson (25);
      std_poisson (25);
(%o55) 25
(%o56) 5

(%i57) pdf_poisson (25, 25), numer;
(%o57) 0.079523
```

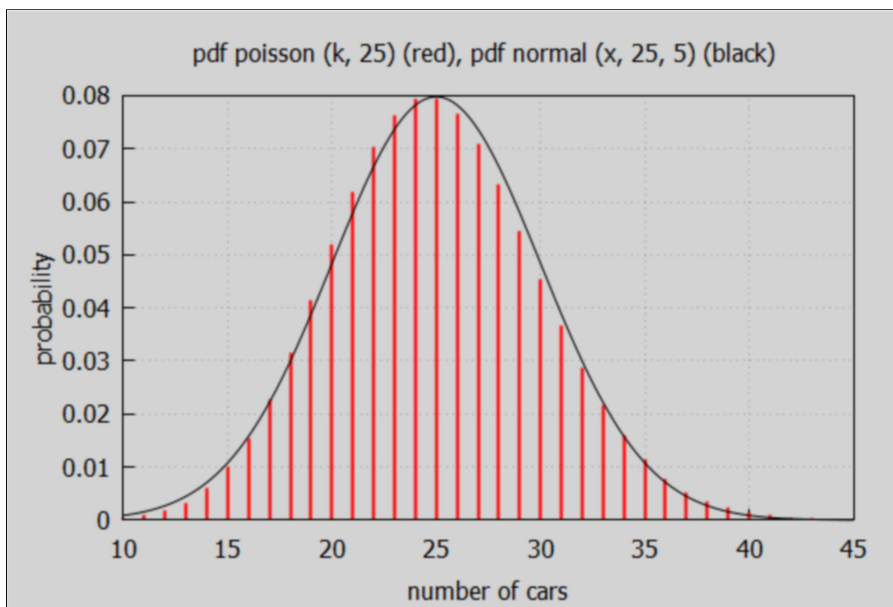
```
(%i61) mypoints : makelist ([k, pdf_poisson (k, 25)], k, 0, 50), numer$
fill (mypoints);
head (mypoints);
tail (mypoints);

(%o59) [[0, 1.3888 10-11], [50, 3.6022 10-6], 51]
(%o60) [[0, 1.3888 10-11], [1, 3.472 10-10], [2, 4.34 10-9]]
(%o61) [[48, 1.412 10-5], [49, 7.2043 10-6], [50, 3.6022 10-6]]
```

We will only show the interval [10, 45] to get a better plot.

```
(%i62) wxdraw2d (xrange = [10, 45], yrange = [0, 0.08], points_joined = impulses,
xlabel = "number of cars", ylabel = "probability",
title = "pdf poisson (k, 25) (red), pdf normal (x, 25, 5) (black) ",
grid = true, color = red, line_width = 2,
background_color = light_gray, points (mypoints),
color = black, line_width = 1,
explicit (pdf_normal (x, 25, 5), x, 10, 45) )$
```

(%t62)



6.6 FS Ex 2: Number of Particles Emitted by Source over Δt

This is Problem 5 on Fred Senese's Ch. 8 discrete pdf worksheet.

A sample of a radioactive isotope emits an average of 27,531 alpha particles over a 1 hour period.

A.) Report the estimated emission rate of alpha particles per second, with an estimated uncertainty.

Let r = number per sec, using 1 hour = 3600 sec.

(%i63) `r : 27531/3600, numer;`

(r) 7.6475

r has units number/sec. The average number for the trial time interval selected (1 sec) is 7.6475, which is λ , and λ is dimensionless.

(%i64) `λ : %;`

(λ) 7.6475

(%i65) `mean_poisson (λ);`

(%o65) 7.6475

The uncertainty is taken to be one standard deviation, produced by `std_poisson (λ)`

(%i66) `std_poisson (λ);`

(%o66) 2.7654

Rounding off both the mean and the standard deviation to two significant figures, we get (7.6 +/- 2.8) particles emitted in the chosen trial time interval of one second.

B.) What is the probability of observing 10 emitted particles during the chosen trial time interval of one second?

(%i67) `pdf_poisson (10, λ), numer;`

(%o67) 0.089984

The chance of observing exactly 10 emitted particles in each trial having the same chosen time interval of one second is about 9%.

C.) You can be 95% confident that the observed count is less than or equal to what value?

(%i68) `quantile_poisson (0.95, λ);`

(%o68) 12

(%i71) `cdf_poisson (11, λ);`

`cdf_poisson (12, λ);`

`cdf_poisson (13, λ);`

(%o69) 0.91183

(%o70) 0.9517

(%o71) 0.97515

We can be 95% confident that the actual count observed will lie in the interval [0, 12].

6.7 Statology 1: Calls per Hour at a Call Center

This example is from:

<https://www.statology.org/poisson-distribution-real-life-examples/>

"Call centers use the Poisson distribution to model the number of expected calls per hour that they'll receive so they know how many call center reps to keep on staff."

"For example, suppose a given call center receives 10 calls per hour. We can use a Poisson distribution calculator to find the probability that a call center receives 0, 1, 2, 3 ... calls in a given hour:

$$P(X = 0 \text{ calls}) = 0.00005$$

$$P(X = 1 \text{ call}) = 0.00045$$

$$P(X = 2 \text{ calls}) = 0.00227$$

$$P(X = 3 \text{ calls}) = 0.00757$$

And so on."

"This gives call center managers an idea of how many calls they're likely to receive per hour and enables them to manage employee schedules based on the number of expected calls."

The chosen trial time period is one hour, there are, say, 8 trials per day, each of one hour, and the measured average number of calls in the chosen trial time period is 10 calls, each trial (no matter what time of day) is assumed to experience the same average rate of calls (10 calls/hr) and to be independent of the other trials' number of calls.

```
(%i72) mean_poisson (10);
```

```
(%o72) 10
```

To find the probability that a call center receives 0, 1, 2, 3 ... calls in a given hour, we can use pdf_poisson (k, 10):

```
(%i73) for k : 0 thru 5 do print (k, float (pdf_poisson (k, 10)))$
```

```
0 4.54 10-5
```

```
1 4.54 10-4
```

```
2 0.00227
```

```
3 0.0075667
```

```
4 0.018917
```

```
5 0.037833
```

The Maxima functions produce the same probabilities as asserted on the Statology Examples webpage. $P(0) = 0.0000454$, $P(1) = 0.000454$, etc.

6.8 Statology 2: Number of Restaurant Customers per Day

<https://www.statology.org/poisson-distribution-real-life-examples/>

"Restaurants use the Poisson distribution to model the number of expected customers that will arrive at the restaurant per day."

"For example, suppose a given restaurant receives an average of 100 customers per day. We can use the Poisson distribution calculator to find the probability that the restaurant receives more than a certain number of customers:"

" $P(X > 110 \text{ customers}) = 0.14714$

$P(X > 120 \text{ customers}) = 0.02267$

$P(X > 130 \text{ customers}) = 0.00171$

And so on."

"This gives restaurant managers an idea of the likelihood that they'll receive more than a certain number of customers in a given day."

In this problem, the chosen time period is one day, and $\lambda = 100 =$ average number of customers/day

$P(k \leq 110) + P(k > 110) = 1$, so $P(k > 110) = 1 - P(k \leq 110) = 1 - \text{cdf_poisson}(110, 100)$

```
(%i74) 1 - cdf_poisson(110, 100), numer;
```

```
(%o74) 0.14714
```

About a 15% likelihood the restaurant will receive more than 110 customers on any given day.

```
(%i76) fpprintprec : 3$
```

```
for k: 110 thru 130 step 10 do print(k, 1 - float(cdf_poisson(k, 100)))$
```

```
110 0.147
```

```
120 0.0227
```

```
130 0.00171
```

About a 2% likelihood the restaurant will receive more than 120 customers on any given day.

About a 0.2% likelihood the restaurant will receive more than 130 customers in a given day.

A chance of 0.2% means the same as a probability of $0.002 = 2 \cdot 10^{-3} = 2 / (1000)$

= "2 parts in a thousand."

6.9 Statology 3: Number of Website Visitors per Hour

"Website hosting companies use the Poisson distribution to model the number of expected visitors per hour that websites will receive."

"For example, suppose a given website receives an average of 20 visitors per hour. We can use the Poisson distribution calculator to find the probability that the website receives more than a certain number of visitors in a given hour:

$$P(X > 25 \text{ visitors}) = 0.11218$$

$$P(X > 30 \text{ visitors}) = 0.01347$$

$$P(X > 35 \text{ visitors}) = 0.00080$$

And so on."

"This gives hosting companies an idea of how much bandwidth to provide to different websites to ensure that they'll be able to handle a certain number of visitors each hour."

```
(%i77) for k: 25 thru 35 step 5 do print(k, 1 - float (cdf_poisson (k, 20)))$
25 0.112
30 0.0135
35 8.04 10-4
```

6.10 Statology 4: Number of Bankruptcies Filed per Month

"Banks use the Poisson distribution to model the number of expected customer bankruptcies per month."

"For example, suppose a given bank has an average of 3 bankruptcies filed by customers each month. We can use the Poisson distribution calculator to find the probability that the bank receives a specific number of bankruptcy files in a given month:

$$P(X = 0 \text{ bankruptcies}) = 0.04979$$

$$P(X = 1 \text{ bankruptcy}) = 0.14936$$

$$P(X = 2 \text{ bankruptcies}) = 0.22404$$

And so on."

"This gives banks an idea of how much reserve cash to keep on hand in case a certain number of bankruptcies occur in a given month."

With an average $\lambda = 3$ bankruptcies/month, we can find the probability the given bank will receive exactly k bankruptcy files in a given month using `pdf_poisson (k, 3)`.

```
(%i78) for k:0 thru 2 do print (k, float (pdf_poisson (k, 3)) )$
0 0.0498
1 0.149
2 0.224
```

About 5% likelihood the given bank will receive no bankruptcy filing in a given month.

About a 15% likelihood the given bank will receive exactly one bankruptcy filing in a given month.

6.11 Statology 5: Number of Network Failures per Week

"Technology companies use the Poisson distribution to model the number of expected network failures per week."

"For example, suppose a given company experiences an average of 1 network failure per week. We can use the Poisson distribution calculator to find the probability that the company experiences a certain number of network failures in a given week:

$$P(X = 0 \text{ failures}) = 0.36788$$

$$P(X = 1 \text{ failure}) = 0.36788$$

$$P(X = 2 \text{ failures}) = 0.18394$$

And so on."

"This gives the company an idea of how many failures are likely to occur each week."

To find the probability the given company will experience exactly k network failures in a given week we can use `pdf_poisson(k, 1)`.

```
(%i79) for k:0 thru 2 do print (k, float (pdf_poisson (k, 1)) )$
0 0.368
1 0.368
2 0.184
```

About 37% likelihood the given company will experience exactly 0 network failures in a given week.

About 37% likelihood the given company will experience exactly 1 network failures in a given week.

About 18% likelihood the given company will experience exactly 2 network failures in a given week.

6.12 Number of Beam Particles per Pulse

The number of beam particles per pulse is assumed to be Poisson distributed. If it is known that the average number of particles per pulse (in a particle accelerator) is 16, what is the probability that in a given pulse the number of particles will lie in the interval [12, 20]?

```
(%i80) sum (pdf_poisson (k, 16), k, 12, 20), numer;
(%o80) 0.741
```


The probability the number of beam particles in any particular pulse will lie in the interval [12, 20] is 74%.

6.13 [RS] Prob. 3.30

"Past experience shows that 1% of the lightbulbs produced in a plant are defective. Find the probability that more than one bulb is defective in a random sample of 30 bulbs, using a.) the binomial distribution and b.) the Poisson distribution.

a.) The binomial distribution is used to find the probability of k 'successes' $P_b(k; n, p)$ in n trials of the same experiment when (1) there are only two possible and mutually exclusive outcomes, (2) the n trials are independent, and (3) the probability of 'success', p , remains constant in each trial.

Using the binomial distribution, the size of the sample is $n = 30$ bulbs, for each bulb the probability of finding it defective (success) is $p = 0.01$, and we are asked to find $P(k > 1)$, the probability that more than one bulb in the sample is defective.

Since $P(k > 1) + P(k \leq 1) = 1$, we have $P(k > 1) = 1 - P(k \leq 1) = 1 - \text{cdf_binomial}(1, n, p)$.

```
(%i81) 1 - cdf_binomial (1, 30, 0.01);
```

```
(%o81) 0.0361
```

The chance of finding more than one bulb defective in a sample of 30 bulbs for which $p = 0.01$ is the probability of finding any random bulb defective is 3.61%.

b.) "Since $n = 30$ and $p = 0.01$, $n \cdot p = 0.3 < 5$, we can use the Poisson approximation of the binomial distribution. Letting $\lambda = n \cdot p$, we have to find $P(k > 1)$, where k is the number of defective bulbs found.

$P(k > 1) = 1 - P(k \leq 1) = 1 - \text{cdf_poisson}(1, \lambda) = 1 - \text{cdf_poisson}(1, 0.3)$.

```
(%i82) 1 - cdf_poisson (1, 0.3);
```

```
(%o82) 0.0369
```

The Poisson distribution approximation to the binomial distribution says, in the example, that the chance that more than one defective bulb is found (when the average is 0.3) is 3.69%.

"As n becomes larger (ie., a larger sample), [with p fixed], the Poisson approximation becomes even closer to the prediction of the binomial distribution."

6.14 Barlow Problem 3.3

"During a meteor shower, meteors fall at the rate 15.7 per hour. What is the probability of observing less than 5 in a given period of 30 minutes?"

The average rate of meteors/(30 min) is $15.7/2 = 7.85 = \lambda$.

$P(k < 5) = P(0) + P(1) + P(2) + P(3) + P(4) = \text{cdf_poisson}(4, 7.85)$

```
(%i83) cdf_poisson(4, 7.85);
```

```
(%o83) 0.109
```

The chance of observing less than 5 meteors over a random period of 30 minutes is 10.9%.

6.15 Barlow Problem 3.4

Repeat the previous problem, using the Gaussian approximation to the Poisson.

Physicists usually call the continuous Normal distribution the "Gaussian distribution". We discuss the Normal distribution in Stat04-Normal.wmx. $\text{pdf_normal}(x, m, s)$ can sometimes be a useful continuous approximation to a Poisson distribution $\text{pdf_poisson}(k, \lambda)$ if $m = \lambda$ and $s = \sqrt{\lambda}$.

Likewise $\text{cdf_normal}(x, m, s)$ as an approximation to $\text{cdf_poisson}(k, \lambda)$.

Since $\lambda = 7.85$ and $\sqrt{\lambda} = 2.802$,

```
(%i84) cdf_normal(4, 7.85, 2.802), numer;
```

```
(%o84) 0.0847
```

which implies a 8.47% chance rather than 10.9% chance.

6.16 Barlow Problem 3.5

A student is trying to hitch a lift. Cars pass at random intervals, at an average rate of 1 per minute. The chance of a car giving a lift is 1%. What is the chance that the student will still be waiting: (a) after 60 cars have passed? (b) after 1 hour?

a.) The probability each passing car will turn into a lift (success) is $p = 0.01$.

The probability of $k = 0$ successful events after $n = 60$ trials with the probability of success p in a given trial is $\text{pdf_binomial}(0, n, p) = \text{pdf_binomial}(0, 60, 0.01)$.

```
(%i85) pdf_binomial(0, 60, 0.01);
```

```
(%o85) 0.547
```

The chance the student will still be waiting after 60 cars have passed is 54.7%.

For 60 cars, each with a probability $(1 - p) = 0.99$ of not giving a lift, the binomial formula gives a probability of no lift after 60 cars = $P_b(0, 60, 0.01) = (0 - 0.01)^{60} = (0.99)^{60} = 0.547$.

```
(%i86) 0.99^60;
```

```
(%o86) 0.547
```

b.) The average number of cars/hour is 60 cars/hour. The average number of lift-giving cars per hour is $0.01 \cdot 60$ cars/hour = 0.6 cars/hour. $\text{pdf_poisson}(0, 0.6)$ = probability 0 lift-giving cars appear in one hour.

```
(%i87) pdf_poisson (0, 0.6);
```

```
(%o87) 0.549
```

The chance the student will still be waiting for a lift after one hour is 54.9%.

"Please get the important difference between the two situations clear, and understand why the two different formulae are applied. Once you understand this, you are on the way to becoming a statistics expert."

6.17 random_poisson (λ), random_poisson (λ , n)

The Maxima function $\text{random_poisson}(\lambda)$ returns a $\text{Poisson}(\lambda)$ random variate, with $\lambda > 0$. The parameter λ is the requested average (mean) value of the distribution.

```
(%i88) for j thru 5 do print (j, random_poisson (15))$
```

```
1 17
```

```
2 16
```

```
3 22
```

```
4 13
```

```
5 15
```

The Maxima function $\text{random_poisson}(\lambda, m)$ returns a list of m Poisson distributed random integers, where λ is the average (mean) of the distribution.

6.17.1 Random Sample Size $m = 10$ Simulations, $\lambda = 15$

```
(%i89) rsample : random_poisson (15, 10);
```

```
(rsample) [13, 14, 20, 12, 12, 17, 12, 11, 14, 21]
```

6.17.2 Random Sample Size $m = 100$ Simulations, $\lambda = 15$

We expect to get more integer repetitions in the returned list of values if we use $\text{random_poisson}(15, 100)$, which returns a list of 100 integers.

```
(%i93) rsample : random_poisson (15, 100)$
      fill (rsample);
      head (rsample);
      tail (rsample);
(%o91) [7, 12, 100]
(%o92) [7, 12, 13]
(%o93) [7, 15, 12]

(%i94) length (rsample);
(%o94) 100
```

6.17.3 discrete_freq (data)

The Maxima function `discrete_freq(aList)` counts the number of unique discrete "readings" of some instrument recorded in the list `aList` and returns a new list:

[list-of-unique-readings, list-of-frequency-of-each-unique-reading].

The elements of the list 'frequencies' (in the following) corresponds to the elements of the list 'uniqueData', element by element.

```
(%i97) [uniqueData, frequencies] : discrete_freq (rsample)$
      uniqueData;
      frequencies;
(%o96) [7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 24, 25, 26]
(%o97) [2, 1, 5, 8, 10, 15, 11, 10, 8, 9, 7, 6, 2, 3, 1, 1, 1]

(%i99) length (uniqueData);
      length (frequencies);
(%o98) 17
(%o99) 17

(%i100) [first (uniqueData), last (uniqueData)];
(%o100) [7, 26]
```

If we add the integers in the list 'frequencies', we will get 100, since `rsample` has 100 elements, and each of the 100 elements (an integer) will contribute unity to one of the `uniqueData` elements.

```
(%i101) Lsum (frequencies);
(%o101) 100
```

Let `nfrequencies` be a list of "normalized" frequencies (so the sum of the elements of `nfrequencies` equals 1 (we need the sum of the probabilities to equal 1)).

```
(%i103) nfrequencies : frequencies/100.0;
```

```
Lsum (nfrequencies);
```

```
(%o103) [0.02,0.01,0.05,0.08,0.1,0.15,0.11,0.1,0.08,0.09,0.07,0.06,0.02,  
0.03,0.01,0.01,0.01]
```

```
(%o103) 1.0
```

What is the largest element of the list `nfrequencies`?

```
(%i104) lmax (nfrequencies);
```

```
(%o104) 0.15
```

To make a plot of probabilities (vertical axis) versus number of events in time interval (horizontal axis), we let `mypoints` be a list of length `(uniqueData)` sublists, with each sublist containing $[x, y] = [\text{number of events in time interval, the corresponding probability}] = [\text{uniqueData}[j], \text{nfrequencies}[j]]$ based on the $m = 100$ simulations.

```
(%i108) mypoints : makelist ([ uniqueData[j], nfrequencies [j] ], j, 1, length (uniqueData))$  
fill (mypoints);
```

```
head (mypoints);
```

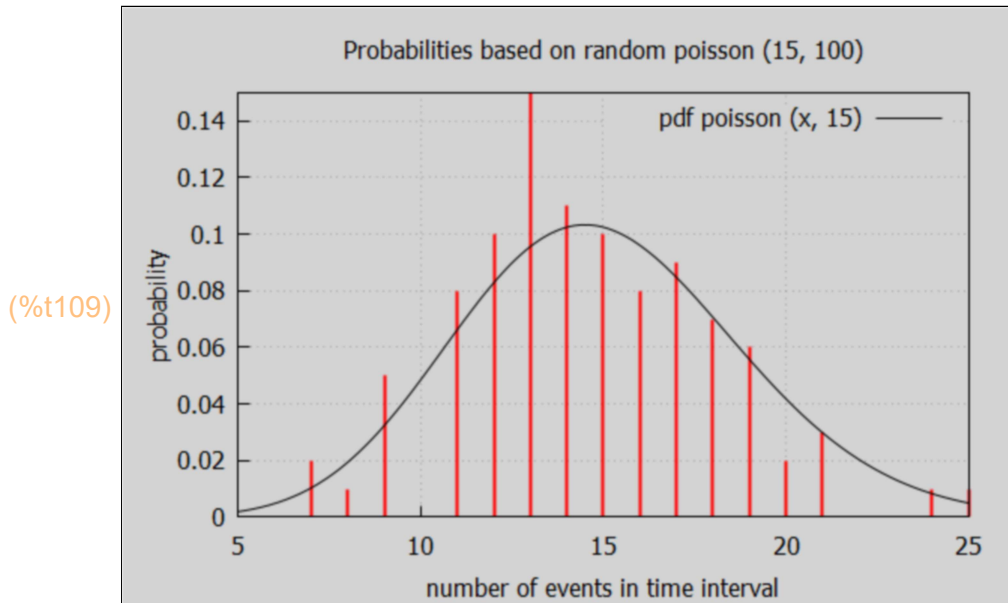
```
tail (mypoints);
```

```
(%o106) [[7,0.02],[26,0.01],17]
```

```
(%o107) [[7,0.02],[8,0.01],[9,0.05]]
```

```
(%o108) [[24,0.01],[25,0.01],[26,0.01]]
```

```
(%i109) wxdraw2d ( xrange = [5, 25], yrange = [0, 0.15], points_joined = impulses,
  xlabel = "number of events in time interval", ylabel = "probability", grid = true,
  title = "Probabilities based on random poisson (15, 100)",
  background_color = light_gray,
  line_width = 2, color = red, points (mypoints),
  color = black, line_width = 1, key = "pdf poisson (x, 15)",
  explicit (pdf_poisson (x, 15), x, 5, 25))$
```



When you run this script on your own, your probabilities (vertical axis) will be slightly different, and you can change the vertical range yrange in draw2d, if needed, to see all your points clearly. You can also change the xrange values based on the first and last elements of uniqueData.

6.17.4 Random Sample Size $m = 1000$ Simulations, $\lambda = 15$

We expect to get more integer repetitions in the returned list of values if we use `random_poisson (15, 1000)`, which returns a list of 1000 integers.

```
(%i113) rsample : random_poisson (15, 1000)$
  fill (rsample);
  head (rsample);
  tail (rsample);
```

```
(%o111) [9,9,1000]
```

```
(%o112) [9,18,13]
```

```
(%o113) [11,10,9]
```

```
(%i116)[uniqueData, frequencies] : discrete_freq (rsample)$
  uniqueData;
  frequencies;

(%o115) [5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,
  26,28,29,31]

(%o116) [1,4,11,14,38,46,68,100,96,99,97,94,83,72,62,31,32,18,13,6
  ,6,5,1,2,1]

(%i118) length (uniqueData);
  length (frequencies);

(%o117) 25
(%o118) 25

(%i119)[first (uniqueData), last (uniqueData)];
(%o119) [5,31]
```

If we add the integers in the list 'frequencies, we will get 1000, since rsample has 1000 elements, and each of the 1000 elements (an integer) will contribute unity to one of the uniqueData elements.

```
(%i120) Lsum (frequencies);
(%o120) 1000
```

Let nfrequencies be a list of "normalized" frequencies (so the sum of the elements of nfrequencies equals 1 (we need the sum of the probabilities to equal 1).

```
(%i121) nfrequencies : frequencies/1000.0;
(nfrequencies) [0.001,0.004,0.011,0.014,0.038,0.046,0.068,0.1,0.096,0.099,
  0.097,0.094,0.083,0.072,0.062,0.031,0.032,0.018,0.013,0.006,0.006,
  0.005,0.001,0.002,0.001]
```

```
(%i122) Lsum (nfrequencies );
(%o122) 1.0
```

What is the largest element of the list nfrequencies?

```
(%i123) lmax (nfrequencies);
(%o123) 0.1
```

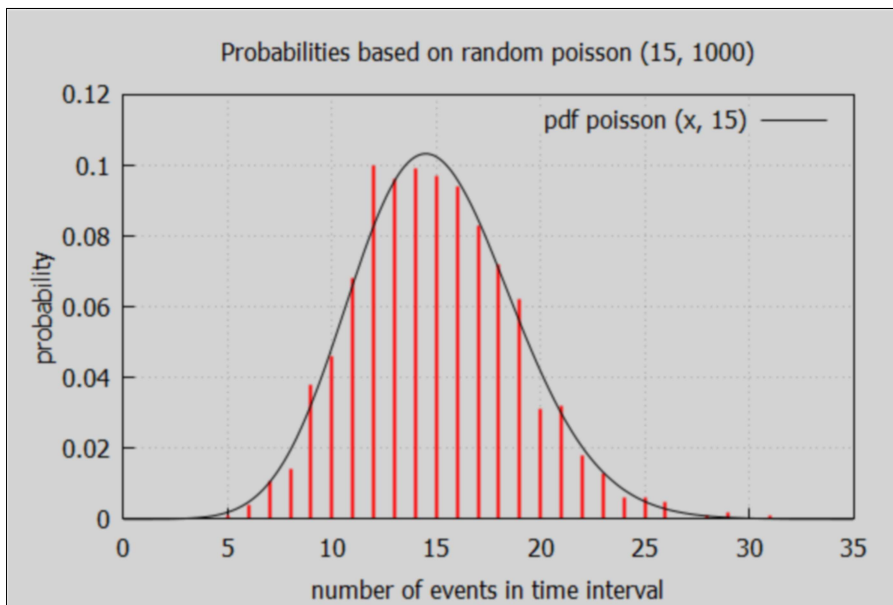
To make a plot of probabilities (vertical axis) versus number of events in the time interval (horizontal axis), we let `mypoints` be a list of length (`uniqueData`) sublists, with each sublist containing $[x, y] = [\text{number of events in the time interval, the corresponding probability}] = [\text{uniqueData}[j], \text{nfrequencies}[j]]$ based on the $m = 100$ simulations.

```
(%i127) mypoints : makelist ([ uniqueData[j], nfrequencies [j] ], j, 1, length (uniqueData))$
fill (mypoints);
head (mypoints);
tail (mypoints);

(%o125) [[5,0.001],[31,0.001],25]
(%o126) [[5,0.001],[6,0.004],[7,0.011]]
(%o127) [[28,0.001],[29,0.002],[31,0.001]]

(%i128) wxdraw2d ( xrange = [0, 35], yrange = [0, 0.12], points_joined = impulses,
xlabel = "number of events in time interval", ylabel = "probability", grid = true,
title = "Probabilities based on random poisson (15, 1000)",
background_color = light_gray,
line_width = 2, color = red, points (mypoints),
color = black, line_width = 1, key = "pdf poisson (x, 15)",
explicit (pdf_poisson (x, 15), x, 0, 35))$
```

(%t128)



With a sample size 1000 we get a closer fit to `pdf_poisson (n, 15)`.