

STA286 Lecture 16

Neil Montgomery

Last edited: 2017-03-02 11:01

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless “dummy” variable):

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless “dummy” variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!}$$

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless “dummy” variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} = e^{-\lambda} e^{e^s \lambda} = e^{\lambda(e^s - 1)}$$

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless “dummy” variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} = e^{-\lambda} e^{e^s \lambda} = e^{\lambda(e^s - 1)}$$

Then it's easy to show $E(X) = \lambda$ and $\text{Var}(X) = \lambda$.

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless “dummy” variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} = e^{-\lambda} e^{e^s \lambda} = e^{\lambda(e^s - 1)}$$

Then it's easy to show $E(X) = \lambda$ and $\text{Var}(X) = \lambda$.

Using full-blown Poisson process notation, we have that the expect value and the variance of the number of events in $[s, t]$ are both $\lambda(t - s)$.

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

The time when the event happened is a continuous random variable, which we can call X . What is its distribution?

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

The time when the event happened is a continuous random variable, which we can call X . What is its distribution?

We can derive its cumulative distribution function from first principles. We know $P(X \leq x | N(t) = 1) = 0$ for $x \leq 0$ and $P(X \leq x | N(t) = 1) = 1$ for $x > t$.

cdf of “the time when that one thing happened” - cont'

Between 0 and t we have:

$$P(X \leq x | N(t) = 1) = 1 - P(X > x | N(t) = 1)$$

cdf of “the time when that one thing happened” - cont'

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1)\end{aligned}$$

cdf of “the time when that one thing happened” - cont'

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1) \\&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)}\end{aligned}$$

cdf of “the time when that one thing happened” - cont'

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1) \\&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)}\end{aligned}$$

cdf of “the time when that one thing happened” - cont’

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1) \\&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)}\end{aligned}$$

cdf of “the time when that one thing happened” - cont’

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1) \\&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{e^{-\lambda x} \lambda(t-x) e^{-\lambda(t-x)}}{\lambda t e^{-\lambda t}}\end{aligned}$$

cdf of “the time when that one thing happened” - cont’

Between 0 and t we have:

$$\begin{aligned}P(X \leq x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\&= 1 - P(N(x) = 0 | N(t) = 1) \\&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)} \\&= 1 - \frac{e^{-\lambda x} \lambda (t - x) e^{-\lambda(t-x)}}{\lambda t e^{-\lambda t}} \\&= 1 - \frac{t - x}{t} = \frac{x}{t}\end{aligned}$$

density of “the time when that one thing happened”

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leq 0 \\ \frac{x}{t} & : 0 < x \leq t \\ 1 & : x > t \end{cases}$$

density of “the time when that one thing happened”

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leq 0 \\ \frac{x}{t} & : 0 < x \leq t \\ 1 & : x > t \end{cases}$$

So the density is:

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} \frac{1}{t} & : 0 < x \leq t \\ 0 & : \text{otherwise} \end{cases}$$

density of “the time when that one thing happened”

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leq 0 \\ \frac{x}{t} & : 0 < x \leq t \\ 1 & : x > t \end{cases}$$

So the density is:

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} \frac{1}{t} & : 0 < x \leq t \\ 0 & : \text{otherwise} \end{cases}$$

The density is flat between 0 and t . We call this a “uniform distribution” between 0 and t .

the uniform distributions

The random variable X , the result of “picking a real number at random between a and b ”, is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{otherwise} \end{cases}$$

the uniform distributions

The random variable X , the result of “picking a real number at random between a and b ”, is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{otherwise} \end{cases}$$

Easy to show $E(X) = (a + b)/2$ and $\text{Var}(X) = (b - a)^2/12$.

the uniform distributions

The random variable X , the result of “picking a real number at random between a and b ”, is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{otherwise} \end{cases}$$

Easy to show $E(X) = (a + b)/2$ and $\text{Var}(X) = (b - a)^2/12$.

mgf is easy to determine, but not really useful for anything.

the uniform distributions

The random variable X , the result of “picking a real number at random between a and b ”, is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{otherwise} \end{cases}$$

Easy to show $E(X) = (a + b)/2$ and $\text{Var}(X) = (b - a)^2/12$.

mgf is easy to determine, but not really useful for anything.

We say $X \sim U[a, b]$, with $U[0, 1]$ an important special case.

Poisson approximation to binomial

The limit $P(X(t) = k) \rightarrow P(N(t) = k)$ converges very fast, which means difficult Binomial calculations can be approximated very accurately using a Poisson probability calculations.

This was great in the 1960s, but not so important now. There is a discussion in the textbook and a handful of textbook exercises we'll call "optional" if you are intested.

waiting time to the first event of a Poisson process

In a Poisson process with rate λ , the first event will happen at some random time X . What is the distribution of X ?

We can derive the cdf from first principles by observing the following:

$$P(X > t) = P(N(t) = 0)$$

waiting time to the first event of a Poisson process

In a Poisson process with rate λ , the first event will happen at some random time X . What is the distribution of X ?

We can derive the cdf from first principles by observing the following:

$$P(X > t) = P(N(t) = 0)$$

The cdf is 0 for $t < 0$. Otherwise:

$$F_X(t) = P(X \leq t) = 1 - P(X > t) = 1 - P(N(t) = 0)$$

waiting time to the first event of a Poisson process

In a Poisson process with rate λ , the first event will happen at some random time X . What is the distribution of X ?

We can derive the cdf from first principles by observing the following:

$$P(X > t) = P(N(t) = 0)$$

The cdf is 0 for $t < 0$. Otherwise:

$$F_X(t) = P(X \leq t) = 1 - P(X > t) = 1 - P(N(t) = 0) = 1 - \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = 1 - e^{-\lambda t}$$

waiting time to the first event of a Poisson process

In a Poisson process with rate λ , the first event will happen at some random time X . What is the distribution of X ?

We can derive the cdf from first principles by observing the following:

$$P(X > t) = P(N(t) = 0)$$

The cdf is 0 for $t < 0$. Otherwise:

$$F_X(t) = P(X \leq t) = 1 - P(X > t) = 1 - P(N(t) = 0) = 1 - \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = 1 - e^{-\lambda t}$$

The density is therefore 0 when $t < 0$ and otherwise:

$$f_X(t) = \lambda e^{-\lambda t}$$

the exponential distributions

We say X has an exponential distribution with rate $\lambda > 0$ when it has density $f_X(t) = \lambda e^{-\lambda t}$ for $t > 0$.

the exponential distributions

We say X has an exponential distribution with rate $\lambda > 0$ when it has density $f_X(t) = \lambda e^{-\lambda t}$ for $t > 0$.

The m.g.f. is:

$$M_X(s) = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda - s}$$

the exponential distributions

We say X has an exponential distribution with rate $\lambda > 0$ when it has density $f_X(t) = \lambda e^{-\lambda t}$ for $t > 0$.

The m.g.f. is:

$$M_X(s) = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda - s}$$

The mean and variance are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$

the exponential distributions

We say X has an exponential distribution with rate $\lambda > 0$ when it has density $f_X(t) = \lambda e^{-\lambda t}$ for $t > 0$.

The m.g.f. is:

$$M_X(s) = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda - s}$$

The mean and variance are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$

Exponential distributions have the *memoryless* property, which is a crucial aspect of Poisson process as a model for “complete randomness”.

the exponential distributions

We say X has an exponential distribution with rate $\lambda > 0$ when it has density $f_X(t) = \lambda e^{-\lambda t}$ for $t > 0$.

The m.g.f. is:

$$M_X(s) = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda - s}$$

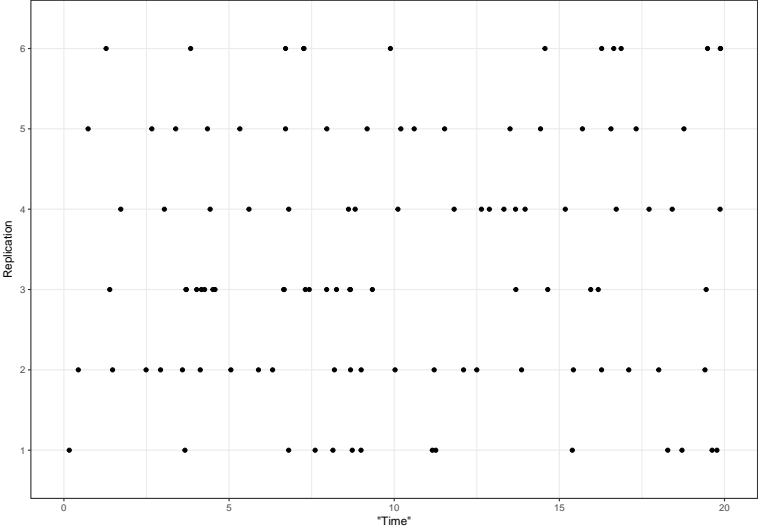
The mean and variance are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$

Exponential distributions have the *memoryless* property, which is a crucial aspect of Poisson process as a model for “complete randomness”.

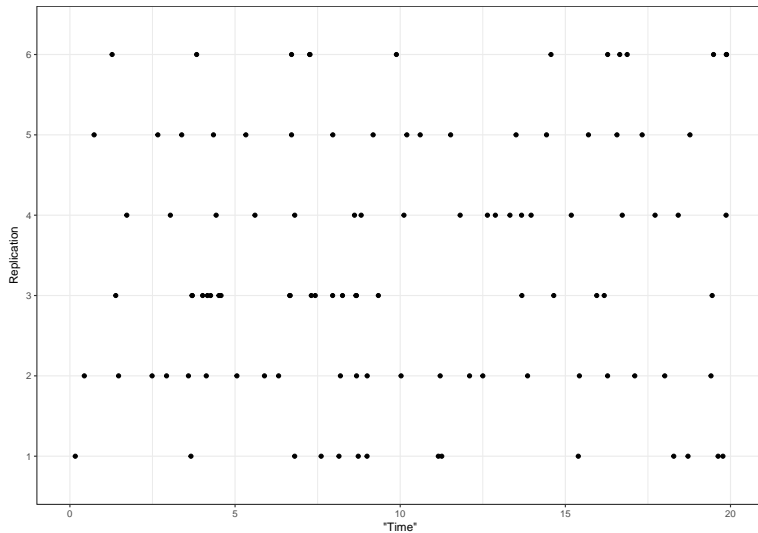
The Exponential distributions are the *only* (continuous) memoryless distributions.

DISASTROUS TREND SHOCKER PANIC HEADLINE

complete randomness is hard for humans



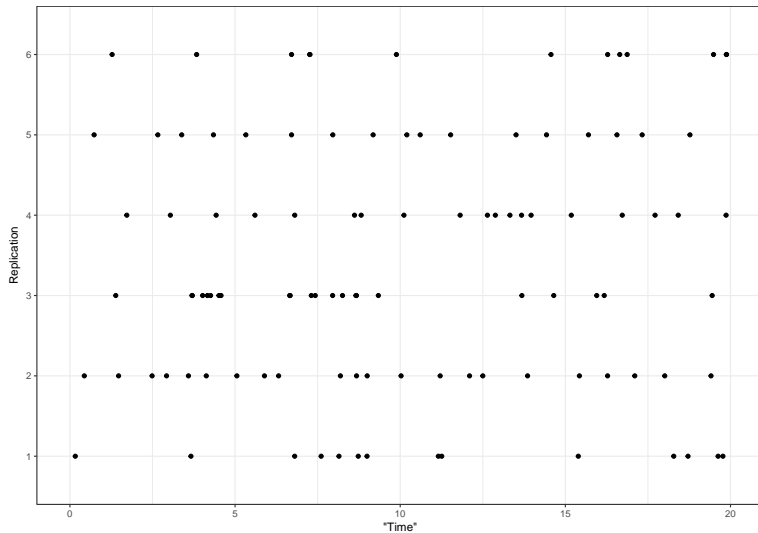
complete randomness is hard for humans



Answer: 1, 3, 6

Most of the time, the number of times a number appears is $\frac{1}{6}$ of the total number of times.

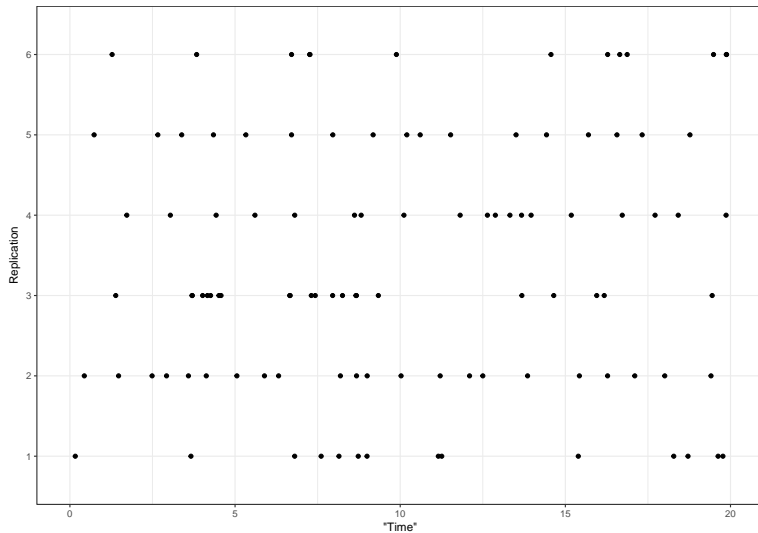
complete randomness is hard for humans



Answer: 1, 3, 6

Most of the time, the distribution of values is not uniform, but follows a pattern that is difficult to predict.

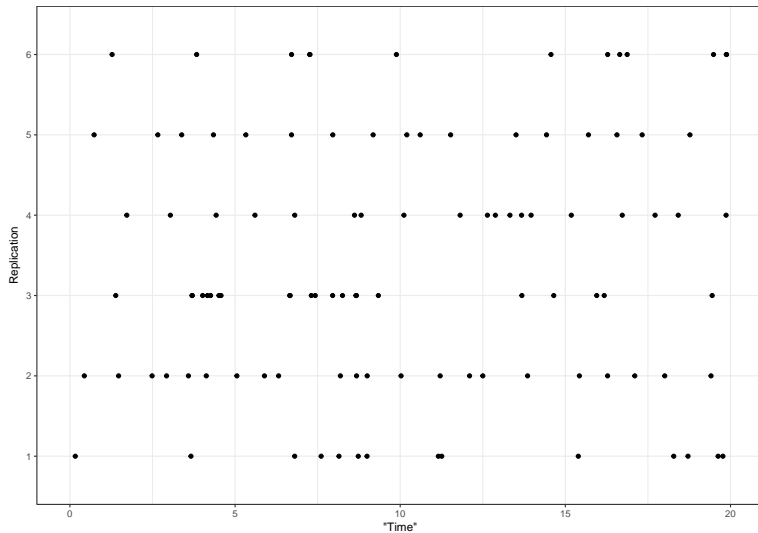
complete randomness is hard for humans



Answer: 1, 3, 6

Most of the time, the distribution of values is not uniform, but follows a pattern that is difficult to predict.

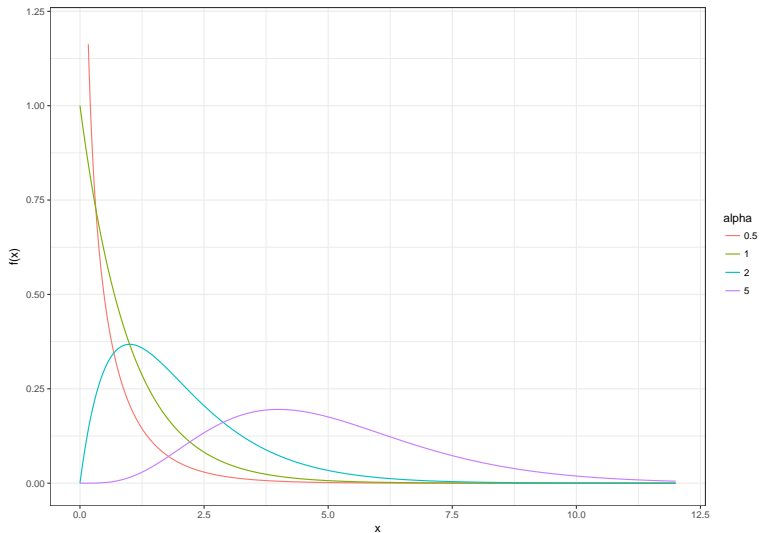
complete randomness is hard for humans



Answer: 1, 3, 6

Most of the time, the distribution of the data is not uniform, with some points appearing more frequently than others, particularly at the extremes of the Y-axis (1 and 6).

pictures of some $\text{Gamma}(\alpha, 1)$ densities



properties of gamma distributions

If $X \sim \text{Gamma}(\alpha, \lambda)$, its moment generating function can be found to be:

$$M_X(s) = \left(\frac{\lambda}{\lambda - s} \right)^\alpha$$

properties of gamma distributions

If $X \sim \text{Gamma}(\alpha, \lambda)$, its moment generating function can be found to be:

$$M_X(s) = \left(\frac{\lambda}{\lambda - s} \right)^\alpha$$

(so the mean and variance are $\frac{\alpha}{\lambda}$ and $\frac{\alpha}{\lambda^2}$)

properties of gamma distributions

If $X \sim \text{Gamma}(\alpha, \lambda)$, its moment generating function can be found to be:

$$M_X(s) = \left(\frac{\lambda}{\lambda - s} \right)^\alpha$$

(so the mean and variance are $\frac{\alpha}{\lambda}$ and $\frac{\alpha}{\lambda^2}$)

Suppose X_1, X_2, \dots, X_n are i.i.d. $\text{Exp}(\lambda)$. What is the distribution of $X = X_1 + X_2 + \dots + X_n$?

properties of gamma distributions

If $X \sim \text{Gamma}(\alpha, \lambda)$, its moment generating function can be found to be:

$$M_X(s) = \left(\frac{\lambda}{\lambda - s} \right)^\alpha$$

(so the mean and variance are $\frac{\alpha}{\lambda}$ and $\frac{\alpha}{\lambda^2}$)

Suppose X_1, X_2, \dots, X_n are i.i.d. $\text{Exp}(\lambda)$. What is the distribution of $X = X_1 + X_2 + \dots + X_n$?

Using the m.g.f. argument it is clear that $X \sim \text{Gamma}(n, \lambda)$, which makes sense in the Poisson process context.