

STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-10-12 14:53

binomial to Poisson

From last time:

- Take an interval $(0, t)$ and an $np = \lambda t$;
- Split into subintervals;
- Stick a Bernoulli trial in each subinterval;
- In such a way that you'd expect about the same number λt of events to occur in $(0, t)$.
- Observe that only one event can happen in each subinterval and non-overlapping collections of subintervals are independent.

Add for today:

- Rather than $(0, t)$ we could have taken any arbitrary (s, t) (different t !). We'll expect about $\lambda(t - s)$ events to take place within (s, t) .

pass to the limit

Define $\lambda(t-s)$ be fixed and always equal to np . This implies $p = \frac{\lambda(t-s)}{n}$.

What happens to $X_n \sim \text{Binomial}\left(n, \frac{\lambda(t-s)}{n}\right)$ distributions as $n \rightarrow \infty$?

$$\lim_{n \rightarrow \infty} P(X_n = k) = \frac{[\lambda(t-s)]^k}{k!} e^{-\lambda(t-s)}$$

Note: λ itself is the rate of occurrences per unit time.

Let $Y \sim \text{Poisson}(\lambda(t-s))$. The above implies more:

$$\lim_{n \rightarrow \infty} P(X_n \in A) = P(Y \in A),$$

for $A \in \mathbb{R}$. So the "distributions" converge.

more on the limit result

This limit result serves a few purposes:

1. Because of the speed and accuracy of the convergence, one can approximate binomial probabilities.
2. Motivate use of Poisson distribution as a suitable probability model (large n , small p , one-at-a-time events, etc.)

binomial approximation example

Suppose $X \sim \text{Binomial}(n, p)$ with $n = 10000$ and $p = 0.001$.

Calculate and approximate $P(X \leq k)$ for $k \in \{0, 1, 2, 3, 4, 5\}$.

Calculation: $\sum_{i=0}^k \binom{10000}{i} (0.001)^i (0.999)^{10000-i}$

Approximation: $\sum_{i=0}^k \frac{10^i e^{-10}}{i!}$, using $\lambda = np = 10$

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
Binom	4.517E-05	4.522E-04	2.263E-03	7.549E-03	1.889E-02	3.780E-02
Poisson	4.540E-05	4.540E-04	2.270E-03	7.567E-03	1.892E-02	3.783E-02
Diff	-2.266E-07	-1.814E-06	-7.031E-06	-1.740E-05	-3.027E-05	-3.785E-05

the "completely random" nature of the Bernoulli process

Only one "event" can happen at a time.

Non-overlapping sets of trials are independent.

Memoryless (as seen via Geometric distributions).

Another idea: suppose by the n^{th} trial exactly 1 event has occurred. Let Y be the index of the trial where the event occurred. This is a random variable. What is its distribution?

Note: sometimes Y is said to have a "discrete uniform distribution on $\{1, 2, \dots, n\}$ "

the Poisson process - I

Suppose as *time* unfolds, events occur, and we keep track of the number of events that occur over time. Denote by $N(t)$ the number of events that happen inside $(0, t]$.

(So $N(t) - N(s)$ is the number of events that happen in some $(s, t]$ with $s < t$).

Let's say we want the occurrence of events to be "completely random" in analogy to a Bernoulli process. Let's specify the following:

- Only one event at a time.
- Counts of events that happen in non-overlapping intervals are independent.

the Poisson process - II

Then if we add a constant "rate" λ per unit time of occurrences it turns out it *must* be that:

$$P(N(t) - N(s) = k) = \frac{[\lambda(t-s)]^k}{k!} e^{-\lambda(t-s)}.$$

We'll say $N(0) = 0$.

Many generalizations exist. The Poisson processes occupy a central role in probability.

practical examples (stolen from *Schay*)

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.
2. More than two will enter in the first two minutes.
3. More than one will enter in each of the first two minutes.

Why and why not might a Poisson process model be suitable here?

continuous random variables

Most discrete random variables arise out of *counting things*. Sometimes it's just more natural for a random variable to take on "any" real value, such as when we're *measuring things*.

Let X be such a random variable. Main focus as always is on *distribution*, i.e. the collection of $P(X \in A)$ for $A \subset \mathbb{R}$. Let's just worry about intervals and consider things like:

$$P(X \in (a, b]) = P(a < X \leq b) = F(b) - F(a).$$

If there is a ("piecewise") continuous function f such that:

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

when we say X is "(absolutely) continuous" and has f as its *probability density function* (or pdf, or just density).

density functions, meanings, consequences - I

Theorem: A pdf completely characterizes a distribution.

Proof: ...

Some corollaries: the cdf F is continuous, $f \geq 0$, and $\int_{-\infty}^{\infty} f(x) dx = 1$.

Advice: *Always* think of a density as living inside its integral.

Heuristic meaning of $f(x)$ can be:

$$f(x)\Delta x \approx \int_x^{x+\Delta x} f(x) dx = P(X \in (x, x + \Delta x])$$

.

density functions, meanings, consequences - II

If X is continuous, then for all $a \in \mathbb{R}$:

$$P(X = a) = P(X \in (a, a]) = \int_a^a f(x) dx = 0$$

Recall from the beginning "pick a random number in $(0,1)$ " and its associated oddities.

example

Pick a number "uniformly" from $(0, 1)$ and let X simply be that number. We have:

$$F(x) = P(X \leq x) = \begin{cases} 0 & : x < 0 \\ x & : 0 \leq x < 1 \\ 1 & : x \geq 1 \end{cases}$$

To find the density, just differentiate:

$$f(x) = F'(x) = \begin{cases} 1 & : 0 \leq x < 1 \\ 0 & : \text{otherwise} \end{cases}$$

The endpoints don't matter. We say X has a uniform distribution with parameters 0 and 1, or $X \sim \text{Uniform}[0, 1]$.

the Uniform[a, b] distributions

Nothing special about 0 and 1. Pick a number between a and b and call it X . The density and cdf will be:

$$f(x) = \begin{cases} \frac{1}{b-a} & : x \in [a, b] \\ 0 & : \text{otherwise;} \end{cases}$$

$$F(x) = \begin{cases} 0 & : x < a \\ \frac{x-a}{b-a} & : x \in [a, b] \\ 1 & : x > b. \end{cases}$$

time to first event of a Poisson process

Let's say we have a Poisson process $N(t)$ with rate λ . The time of the first event is random. Call this time X .

What can we say about X ? Can we completely describe its distribution? Yes!

We can show:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x} & : x > 0 \\ 0 & : \text{otherwise} \end{cases}.$$

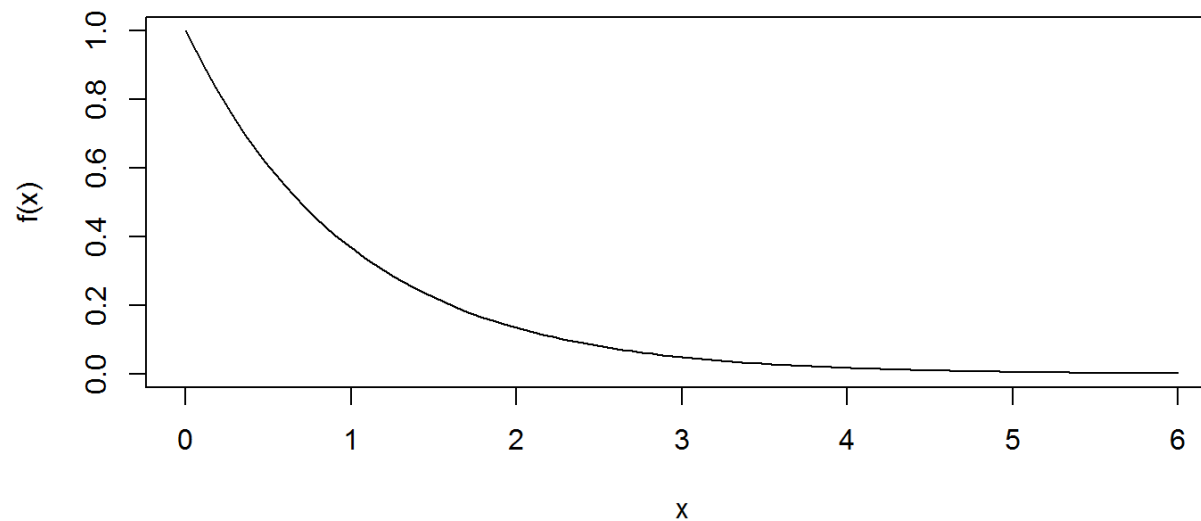
So the density is:

$$f(x) = F'(x) = \begin{cases} \lambda e^{-\lambda x} & : x > 0 \\ 0 & : \text{otherwise.} \end{cases}$$

the exponential distributions

In this case we say X has an exponential distribution with (rate) parameter λ , or $X \sim \text{Exp}(\lambda)$.

Free picture for $\text{Exp}(1)$ density:



what should we expect of a Poisson waiting time?

The Poisson process is the continuous time analogy of the Bernoulli process.
Both "completely random".

It turns out the exponential distributions are also "memoryless", and the proof is pretty much the same.

Theorem: what I just wrote.

Proof: ...

Not only that, but exponential distributions are the *only* memoryless continuous distributions.

Theorem: what I just wrote.

Proof: ...

the hazard function - I

We're always concerned with characterizing distributions. So far we can do this with these functions:

- cdf $F(x) = P(X \leq x)$ (always)
- pmf $p(x) = P(X = x)$ (X discrete only)
- pdf $f(x) = F'(x)$ (X continuous only)
- "survival"/"reliability" function $S(x) = R(x) = P(X > x) = 1 - P(X \leq x) = 1 - F(x)$

Let's add another, based on the idea of "instantaneous probability of occurrence".

Think in terms of a random variable X that measures a time-to-event, and let's say the event *did not happen* by time x . What's the chance the event then happens "right away"?

the hazard function - II

If X is a continuous random variable taking on only positive values, with pdf f and survival function $S(x) = 1 - F(x)$, its *hazard function* h is defined as:

$$h(x) = \frac{f(x)}{S(x)}$$

As a key example, consider $X \sim \text{Exponential}(\lambda)$.

X is "memoryless". What, then, ought to happen with its hazard function?

