Lecture 5: More distributions

MVE055 / MSG810
Mathematical statistics and discrete mathematics

Moritz Schauer

Last updated September 8, 2021, 2021

GU & Chalmers University of Technology

We have seen

- Probability mass functions $f P(X = k) = (k)$.

We have seen

- Probability mass functions $f P(X = k) = (k)$.
- Bernoulli – $\mathrm{Bernoulli}(p)$: $X \in \{0, 1\}$

We have seen

- Probability mass functions $f P(X = k) = (k)$.
- Bernoulli – $\mathrm{Bernoulli}(p)$: $X \in \{0, 1\}$
- Binomial – $\mathrm{Bin}(n, p)$: $X \in \{0, 1, \ldots, n\}$

We have seen

- Probability mass functions $fP(X = k) = (k)$.
- Bernoulli – $\mathrm{Bernoulli}(p)$: $X \in \{0, 1\}$
- Binomial – $\mathrm{Bin}(n, p)$: $X \in \{0, 1, \ldots, n\}$
- Geometric – $\mathrm{Geom}(p)$: $X \in \{1, 2, 3, 4, \ldots\}$

We have seen

- Probability mass functions $fP(X = k) = (k)$.
- Bernoulli – $\mathrm{Bernoulli}(p)$: $X \in \{0, 1\}$
- Binomial – $\mathrm{Bin}(n, p)$: $X \in \{0, 1, \dots, n\}$
- Geometric – $\mathrm{Geom}(p)$: $X \in \{1, 2, 3, 4, \dots\}$
- Normal – $\mathrm{N}(\mu, \sigma^2)$: $X \in (-\infty, \infty)$

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?

$\mathsf{E}(X) = \boxed{}$. $\mathrm{Var}(X) = $.

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?
$\mathsf{E}(X) = np$. $\qquad \mathrm{Var}(X) = \qquad\qquad$.

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?

$\mathsf{E}(X) = np$. $\mathrm{Var}(X) = \boxed{}$.

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?
$\mathsf{E}(X) = np$. $\quad\quad \mathrm{Var}(X) = np(1 - p)$.

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?
$\mathsf{E}(X) = np$. $\qquad \mathrm{Var}(X) = np(1-p)$.

What was the mean and the variance of $X \sim \mathrm{Bin}(n, p)$?
$\mathsf{E}(X) = np$. $\qquad \mathrm{Var}(X) = np(1-p)$.

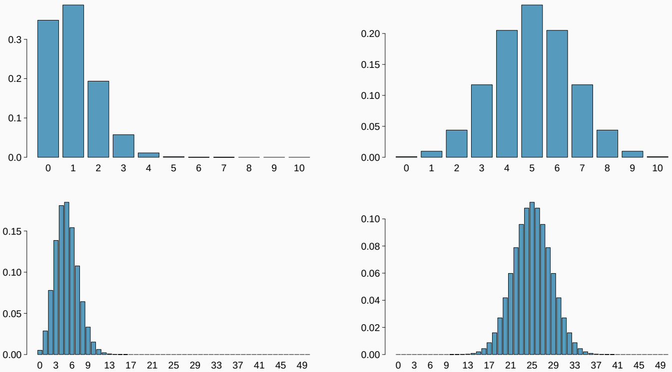
Normal approximation of Binomial distribution

If $X \sim \mathrm{Bin}(n, p)$, X is approximately normally distributed with mean np and variance $np(1-p)$,

$$X \overset{\text{approx.}}{\sim} \mathrm{N}(np, np(1-p)),$$

if both $np > 5$ and $n(1-p) > 5$.

Normal approximation



$n = 10$

$n = 50$

$p = 0.1$ $p = 0.5$

Discrete distributions today

- Poisson distribution – $\text{Poisson}(\mu)$: model the number of events that occur in a time interval, in a region or in some volume.

Discrete distributions today

- Poisson distribution – $\mathrm{Poisson}(\mu)$: model the number of events that occur in a time interval, in a region or in some volume.
- Negative binomial distribution – $\mathrm{nBin}(r, p)$: The number of trials X in a sequence of independent $\mathrm{Bernoulli}(p)$ trials before r successes occur

Discrete distributions today

- Poisson distribution – $\mathrm{Poisson}(\mu)$: model the number of events that occur in a time interval, in a region or in some volume.
- Negative binomial distribution – $\mathrm{nBin}(r, p)$: The number of trials X in a sequence of independent $\mathrm{Bernoulli}(p)$ trials before r successes occur
- Hypergeometric distribution – $\mathrm{Hyp}(N, n, r)$: Draw sample of n objects without replacement out of N. The random variable X is the number of marked objects.

Poisson distribution

The Poisson distribution is often used to model the number of events that occur in a time interval, in a region or in some volume.

(Named after Simeon Denis Poisson, 1781-1840.)

Poisson distribution

The Poisson distribution is often used to model the number of events that occur in a time interval, in a region or in some volume.

(Named after Simeon Denis Poisson, 1781-1840.)

Some examples where this distribution fits well are

- The number of particles emitted per minute (hour, day) of a radioactive material.
- Call connections routed via a cell tower (GSM base station).

Poisson distribution

$$X \sim \text{Poisson}(\mu)$$

A random variable X has Poisson distribution with parameter μ if

$$\mathrm{P}(X = k) = \frac{\mathrm{e}^{-\mu}\mu^k}{k!}, \quad k \in \{0, 1, 2, \dots\}.$$

Poisson distribution

$$X \sim \text{Poisson}(\mu)$$

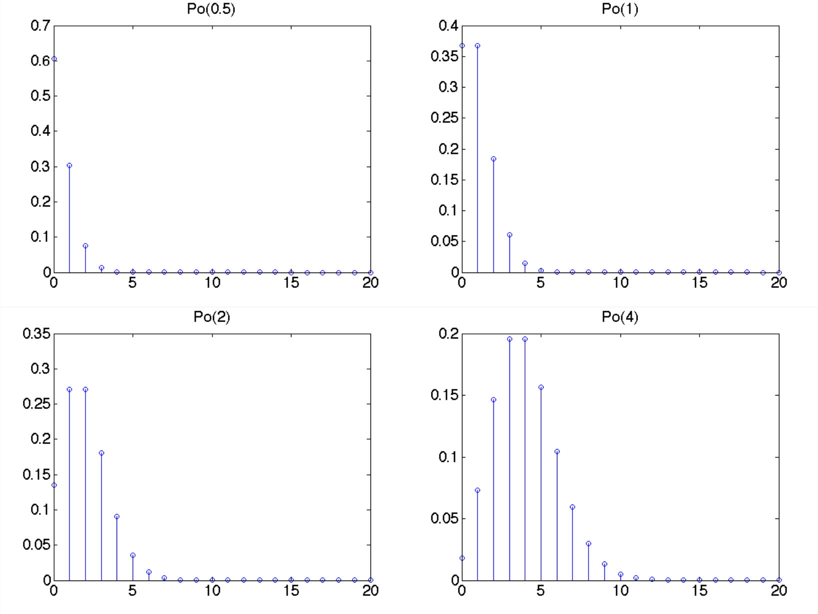
A random variable X has Poisson distribution with parameter μ if

$$\mathrm{P}(X = k) = \frac{\mathrm{e}^{-\mu}\mu^k}{k!}, \quad k \in \{0, 1, 2, \ldots\}.$$

Sum of Poisson distributed random variables.

If $X_1 \sim \text{Poisson}(\mu_1)$ and $X_2 \sim \text{Poisson}(\mu_2)$ are independent, then $X_1 + X_2 \sim \text{Poisson}(\mu_1 + \mu_2)$.

Poisson distribution



Number of chewing gums on a tile is approximately Poisson.

Example

Let X be the number of typos on a printed page with a mean of 3 typos per page. Assume the typos occur independently of each other.

1. What is the probability that a randomly selected page has at least one typo on it?

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

Example

Let X be the number of typos on a printed page with a mean of 3 typos per page. Assume the typos occur independently of each other.

1. What is the probability that a randomly selected page has at least one typo on it?

$$\mathrm{P}(X \geqslant 1) = 1 - \mathrm{P}(X = 0) = 1 - f(0) = 1 - e^{-3}$$

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

Example

Let X be the number of typos on a printed page with a mean of 3 typos per page. Assume the typos occur independently of each other.

1. What is the probability that a randomly selected page has at least one typo on it?

$\mathrm{P}(X \geqslant 1) = 1 - \mathrm{P}(X = 0) = 1 - f(0) = 1 - e^{-3}$

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

Example

Let X be the number of typos on a printed page with a mean of 3 typos per page. Assume the typos occur independently of each other.

1. What is the probability that a randomly selected page has at least one typo on it?

$\mathrm{P}(X \geqslant 1) = 1 - \mathrm{P}(X = 0) = 1 - f(0) = 1 - e^{-3}$

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

In this case $\lambda = 9$ since we have in average 9 typos on three printed pages.

Example

Let X be the number of typos on a printed page with a mean of 3 typos per page. Assume the typos occur independently of each other.

1. What is the probability that a randomly selected page has at least one typo on it?

$\mathrm{P}(X \geqslant 1) = 1 - \mathrm{P}(X = 0) = 1 - f(0) = 1 - e^{-3}$

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

In this case $\lambda = 9$ since we have in average 9 typos on three printed pages.

$\mathrm{P}(X > 8) = 1 - \mathrm{P}(X \leqslant 8) \approx 1 - 0.456$ by table II page 692

Poisson distribution as limit of a Binomial distribution

The Poisson distribution appears as limit of the Binomial distribution if n becomes large and p goes to 0:

Theorem

Let $n \to \infty$, $p \to 0$, and also $np \to \mu$. Then for fix $k \geqslant 0$

$$\binom{n}{k} p^k (1-p)^{n-k} \to \frac{\mu^k e^{-\mu}}{k!} \tag{0.1}$$

Poisson distribution as limit of a Binomial distribution

The Poisson distribution appears as limit of the Binomial distribution if n becomes large and p goes to 0:

Theorem

Let $n \to \infty$, $p \to 0$, and also $np \to \mu$. Then for fix $k \geqslant 0$

$$\binom{n}{k} p^k (1-p)^{n-k} \to \frac{\mu^k e^{-\mu}}{k!} \tag{0.1}$$

Connection to the previous example:

- There is a large number n of atoms in the material and the probability that an atom decays in a unit of time p is very small.

Negative binomial distribution

The number of trials X in a sequence of independent Bernoulli(p) trials before r successes occur has the negative binomial distribution.

Negative binomial distribution

$$X \sim \mathrm{nBin}(r, p)$$

The random variable X has a negative binomial distribution with parameter r and p if

$$\mathrm{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1 \ldots$$

Negative binomial distribution

$$X \sim \mathrm{nBin}(r, p)$$

The random variable X has a negative binomial distribution with parameter r and p if

$$\mathrm{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1 \ldots$$

Motivation: Probability of r successes in k trials: $(1-p)^{k-r} p^r$. The last attempt succeeds. The binomial coefficient gives the number of ways we assign the remaining $r-1$ successes to the remaining $k-1$ trials.

Hypergeometric distribution

- Suppose we have N objects of which r are "marked".
- Draw sample of n objects without replacement. The random variable X is the number of marked objects. Then X has hypergeometric distribution with parameters N, n, r.

Hypergeometric distribution

$$X \sim \mathrm{Hyp}(N, n, r)$$

The random variable X has hypergeometric distribution with parameters N, n and r if

$$\mathrm{P}(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \quad \max(0, n + r - N) \leqslant k \leqslant \min(n, r)$$

Hypergeometric distribution

$$X \sim \mathrm{Hyp}(N, n, r)$$

The random variable X has hypergeometric distribution with parameters N, n and r if

$$\mathrm{P}(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \quad \max(0, n + r - N) \leqslant k \leqslant \min(n, r)$$

If $n = 1$ then $\mathrm{Hyp}(N, 1, r) = \mathrm{Bernoulli}(r/N)$. If N and r are large compared to n we have $\mathrm{Hyp}(N, n, r) \approx \mathrm{Bin}(n, r/N)$.

Continuous distributions today (all positive)

- Exponential distribution – $\mathrm{Exp}(\lambda)$: Time between calls/visitors/people knocking on your door. (Poisson: How many ticks. Exponential: time between ticks.)

Continuous distributions today (all positive)

- Exponential distribution – $\mathrm{Exp}(\lambda)$: Time between calls/visitors/people knocking on your door. (Poisson: How many ticks. Exponential: time between ticks.)
- Gamma distribution – $\Gamma(\alpha, \beta)$: Flexible distribution for positive random variables.

Continuous distributions today (all positive)

- Exponential distribution – $\mathrm{Exp}(\lambda)$: Time between calls/visitors/people knocking on your door. (Poisson: How many ticks. Exponential: time between ticks.)
- Gamma distribution – $\Gamma(\alpha, \beta)$: Flexible distribution for positive random variables.
- χ^2-distribution – $\chi^2(n)$: Distribution for sum of squares of n independent $N(0, 1)$ random variables.

Exponential distribution

$$X \sim \mathrm{Exp}(\lambda)$$

The density function of an exponential distribution with rate λ or is given by

$$f(x) = \lambda \mathrm{e}^{-\lambda x}, \quad x \geqslant 0$$

or equivalently $f(x) = \frac{1}{\beta} e^{-x/\beta}$ where $\beta = \frac{1}{\lambda}$ is the scale.

$\mathsf{E}[X] = \beta$ and $\mathrm{Var}(X) = \beta^2$

The cumulative distribution function is given by

$$F(x) = 1 - e^{-\lambda x}.$$

Exponential distribution

Assume objects arrive after exponentially distributed interarrival times.

λ - how many arrivals per time unit.
β - expected waiting time

Gamma distribution

$$X \sim \mathrm{Gamma}(\alpha, \beta)$$

A random variable X with density function

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x > 0$$

for $\beta > 0$ and $\alpha > 0$ has a Gamma distribution with parameters shape α and scale β, or .

$\mathsf{E}[X] = \alpha\beta$ and $\mathrm{Var}(X) = \alpha\beta^2$.

If X follows a Gamma distribution with parameters α and β, then the m.g.f is given by $m_X(t) = (1 - \beta t)^{-\alpha}$.

χ^2-distribution

$$X \sim \chi^2(n)$$

The Gamma distribution with parameters $\beta = 2$ and $\alpha = \frac{n}{2}$ is called χ^2 -distribution with n degrees of freedom.

$\mathsf{E}[X] = n$ and $\mathrm{Var}(X) = 2n$.

Sum of squares

If $Z_1, \ldots, Z_n$ have standard normal distributions and are independent, then $Z_1^2 + \cdots + Z_n^2$ follow a χ^2 -distribution with n degrees of freedom.

Moment generating function (m.g.f.)

Let X be a random variable

- The k^{th} moment for X is defined by $\mathsf{E}(X^k)$.

Moment generating function (m.g.f.)

Let X be a random variable

- The k^{th} moment for X is defined by $\mathsf{E}(X^k)$.
- The moment generating function for X is defined by

$$m_X(t) = \mathsf{E}(\mathrm{e}^{tX}).$$

Moment generating function (m.g.f.)

Let X be a random variable

- The k^{th} moment for X is defined by $\mathsf{E}(X^k)$.
- The moment generating function for X is defined by

$$m_X(t) = \mathsf{E}(\mathrm{e}^{tX}).$$

- Let $m_X(t)$ be the m.g.f for X. Then

$$\left.\frac{\mathrm{d}^k m_X(t)}{\mathrm{d}t^k}\right|_{t=0} = \mathsf{E}\left(X^k\right)$$

Moment generating function for standard normal distribution

Let $Z \sim \mathsf{N}(0,1)$. Compute the mgf.

Moment generating function for standard normal distribution

Let $Z \sim \mathsf{N}(0, 1)$. Compute the mgf. Use $h(x) = \mathrm{e}^{tx}$ and transform:

Moment generating function for standard normal distribution

Let $Z \sim \mathsf{N}(0,1)$. Compute the mgf. Use $h(x) = \mathrm{e}^{tx}$ and transform:

$$m_X(t) = \mathsf{E}\left[e^{tX}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{\mathrm{e}^{tx}}_{h(x)} \mathrm{e}^{-\frac{1}{2}x^2} \mathrm{d}x$$

Moment generating function for standard normal distribution

Let $Z \sim \mathsf{N}(0,1)$. Compute the mgf. Use $h(x) = \mathrm{e}^{tx}$ and transform:

$$m_X(t) = \mathsf{E}\left[e^{tX}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{\mathrm{e}^{tx}}_{h(x)} \mathrm{e}^{-\frac{1}{2}x^2} \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(x-t)^2} \mathrm{e}^{\frac{1}{2}t^2} \mathrm{d}x = \mathrm{e}^{\frac{1}{2}t^2}$$