Lesson 16: Some Important Continuous RVs

Meike Niederhausen and Nicky Wakim

2025-11-19

Learning Objectives

1. Distinguish between Uniform, Exponential, Gamma, and Normal distributions when reading a word problem.
2. Identify the variable and the parameters in a word problem, and state what the variable and parameters mean.
3. Use the formulas for the pdf/CDF, expected value, and variance to answer questions and find probabilities.

Where are we?

Basics of probability

- Outcomes and events
- Sample space
- Simulations
- Probability axioms
- Probability properties
- Counting
- Rules of probability

Probability for random variables

- Probability functions: pmfs/pdfs
- Cumulative distribution functions (CDFs)
- Important distributions

Two random variables

- Transformations
- Joint distributions
- Independence and conditioning

Expected values

- Means / expected values
- Linear combinations and conditioning
- Variance

Advanced probability

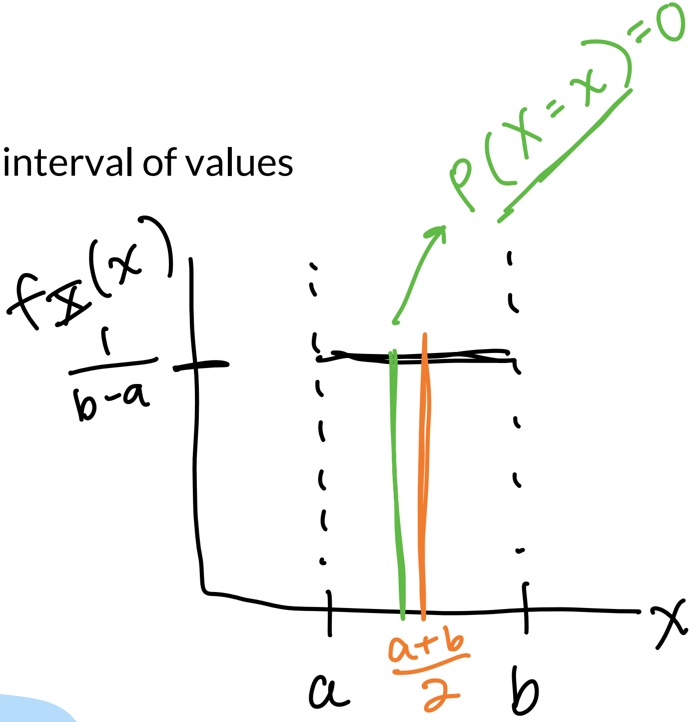
- Central limit theorem
- Functions: moment generating functions

Continuous uniform RVs

Properties of continuous uniform RVs

- **Scenario:** Events are equally likely to happen anywhere or anytime in an interval of values
- Shorthand: $X \sim \mathrm{U}[a, b]$

$$f_X(x) = \frac{1}{b-a}, \text{ for } a \le x \le b$$



$$F_X(x) = \int_a^x f_X(s)\,ds$$

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$

$$\mathrm{E}(X) = \frac{a+b}{2}, \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

$$\mathrm{E}(X) = \int_a^b \underline{x f_X(x)}\,dx$$

Identifying continuous uniform RV from word problems

- Look for some indication that all events are **equally likely**
  - Could also say "uniformly distributed"
- Look for an interval
  - **Time example:** Costumer in your store will approach the cash register in next 30 minutes. Approaching the register throughout the 30 minutes is equally likely.
  - **Length example:** You have a 12 inch string that you need to cut. You are equally likely to cut anywhere on the string.
- Different than the discrete uniform
  - Discrete usually includes a *countable* number of events that are equally likely
  - Continuous is not countable
    - Exact time and length can be measured with infinite decimal places

Helpful R code

$X=0$ $X=60$

Let's say we're looking at equally likely arrival times between 10 am and 11 am.

- If we want to know the probability that someone arrives at 10:30am or earlier:

```
punif(q = 30, min = 0, max = 60)
[1] 0.5
```

$P(X \le 30)$

- If we want to know the time, say t, where the probability of arriving at t or earlier is 0.35:

```
qunif(p = 0.35, min = 0, max = 60)
[1] 21
```

$t=21$

$P(X \le t) = 0.35$

- If we want to know the probability that someone arrives between 10:14 and 10:16 am:

```
punif(q = 16, min = 0, max = 60) - punif(q = 14, min = 0, max = 60)
[1] 0.03333333
```

$P(14 < X < 16) = P(X<16) - P(X<14)$
$= F_X(16) - F_X(14)$

- If we want to sample 20 arrival times from the distribution:

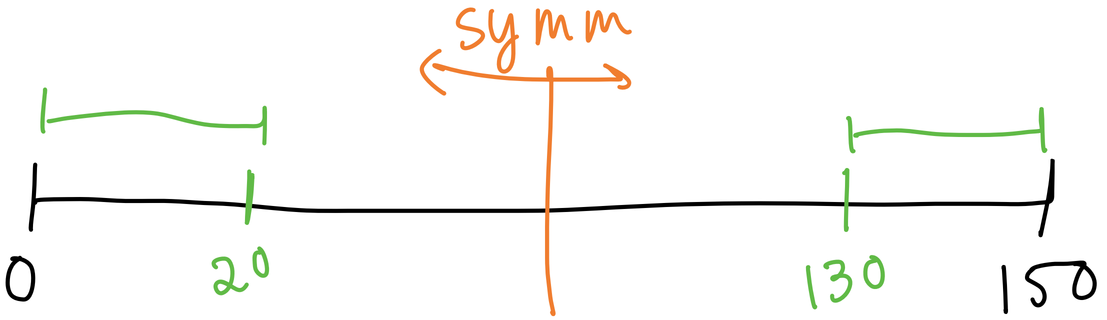
```
runif(n = 20, min = 0, max = 60)
 [1] 57.254766 34.877635  9.463305  9.299211 44.853180 24.713062  7.637811
 [8] 56.432479 57.911681 26.541762 41.756752 32.239442 54.968374  4.029851
[15] 33.138835 40.718886  5.610676 19.739926 21.084020 36.677654
```

$P(X=x) = 0$

Bird on a wire (TB 31.5)

Example 1

A bird lands at a location that is Uniformly distributed along an electrical wire of length 150 feet. The wire is stretched tightly between two poles. What is the probability that the bird is 20 feet or less from one or the other of the poles?



$$P(X<20) + P(X>130)$$

$$2 \cdot P(X<20)$$

$$2 \cdot P(X>130)$$

$$= 2 \cdot \frac{1}{150} \cdot 20 = 0.2667$$

Exponential RVs

Properties of exponential RVs

- **Scenario:** Modeling the time until the next (first) event
- Continuous analog to the geometric distribution!
- Shorthand: $X \sim \mathrm{Exp}(\lambda)$

$\lambda =$ avg rate of success/events

$\lambda =$ avg # of events/success per time/period

$$f_X(x) = \lambda e^{-\lambda x} \text{ for } x > 0, \lambda > 0$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \underline{1 - e^{-\lambda x}} & x \geq 0 \end{cases}$$

$$\mathrm{E}(X) = \frac{1}{\lambda}$$

$$\mathrm{Var}(X) = \frac{1}{\lambda^2}$$

Memoryless Property

If $b > 0$,

$$P(X > a + b | X > a) = P(X > b)$$

- **This can be interpreted as:**
  - If you have waited a seconds (or any other measure of time) without a success
  - Then the probability that you have to wait b more seconds is the same as ~~as~~ the probability of waiting b seconds initially.

Identifying exponential RV from word problems

- Look for time between events/successes
- Look for a rate of the events over time period
- How does it differ from the geometric distribution?
  - Geometric is *number of trials* until first success
  - Exponential is *time* until first success
- Relation to the Poisson distribution?
  - When the time between arrivals is exponential, the number of arrivals in a fixed time interval is Poisson with the mean λ

Helpful R code

Let's say we're sitting at the bus stop, measuring the time until our bus arrives. We know the bus comes every 10 minutes on average.

$\lambda = 1 \text{ bus}/10 \text{ min}$

- If we want to know the probability that the bus arrives in the next 5 minutes:

```
pexp(q = 5, rate = 1/10)
[1] 0.3934693
```

$P(X \leq 5)$

- If we want to know the time, say t, where the probability of the bus arriving at t or earlier is 0.35:

```
qexp(p = 0.35, rate = 1/10)
[1] 4.307829
```

$P(X \leq t) = 0.35$ inverse CDF $F^{-1}(x)$

- If we want to know the probability that the bus arrives between 3 and 5 minutes:

```
pexp(q = 5, rate = 1/10) - pexp(q = 3, rate = 1/10)
[1] 0.1342876
```

- If we want to sample 20 bus arrival times from the distribution:

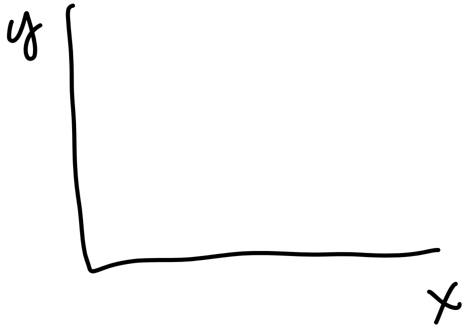
```
rexp(n = 20, rate = 1/10)
 [1]  3.27079121  9.21887422 23.15584516  1.43197788  7.13709980  2.46155182
 [7]  1.25854357 16.49368095  2.07284506  9.61934986  1.72675516  9.66994375
[13]  0.60112111  8.14090359 27.93188929  0.02541741 22.76283795  4.51165567
[19] 15.79182468  2.31727717
```

Transformation of independent exponential RVs

~~Revisit after joint notes~~ min:

Example 2 CDF method

Let $X_i \sim \mathrm{Exp}(\lambda_i)$ be independent RVs, for $i = 1 \ldots n$. Find the pdf for the first of the arrival times.



Let M = first of the arrival times

$= \min(X_1, X_2, X_3, \ldots, X_n)$

$X_i \sim \mathrm{Exp}(\lambda_i)$

$$F_M(m) = P(M \le m) = P(\min(X_1, X_2, \ldots, X_n) \le m)$$

$$= 1 - P(\min(X_1, X_2, \ldots X_n) > m)$$

$$= 1 - P(X_1 > m, X_2 > m, X_3 > m, \ldots X_n > m)$$

$$= 1 - P(X_1 > m)\,P(X_2 > m) \cdots P(X_n > m)$$

$$= 1 - [1 - F_{X_1}(m)][1 - F_{X_2}(m)] \cdots [1 - F_{X_n}(m)]$$

$$= 1 - [\cancel{1} - (\cancel{1} + e^{-\lambda_1 m})][\cancel{1} - (\cancel{1} + e^{-\lambda_2 m})] \cdots [\cancel{1} - (\cancel{1} + e^{-\lambda_n m})]$$

$$= 1 - (e^{-\lambda_1 m} \cdot e^{-\lambda_2 m} \cdots e^{-\lambda_n m})$$

Gamma RVs

$$= 1 - e^{-\sum_{i=1}^{n} \lambda_i m}$$

$$= 1 - e^{-m \sum_{i=1}^{n} \lambda_i}$$

$$e^a e^b = e^{a+b}$$

$$f_M(m) = \frac{d}{dm}\left(1 - e^{-m \sum_{i=1}^{n} \lambda_i}\right) = +\left(\cancel{-}\sum_{i=1}^{n} \lambda_i\right) e^{-m \sum_{i=1}^{n} \lambda_i}$$

$$\frac{d}{dx} e^{ax} = a e^{ax}$$

$$\hookrightarrow f_M(m) = \left(\sum_{i=1}^{n} \lambda_i\right) e^{-m \sum_{i=1}^{n} \lambda_i}$$

for $m > 0$

$\lambda_i >$

$$M = \min(X_1, X_2, \dots X_n)$$

$$\sim \mathrm{Exp}\left(\lambda = \sum_{i=1}^{n} \lambda_i\right)$$

when X_i's independent

Properties of gamma RVs

- **Scenario:** Modeling the time until the r^{th} event.
- Continuous analog to the Negative Binomial distribution
- Shorthand: $X \sim \text{Gamma}(r, \lambda)$ or $X \sim \text{Gamma}(\alpha, \beta)$

gamma

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \text{ for } x > 0, \lambda > 0, \boxed{\Gamma(r) = (r-1)!}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} \sum_{j=0}^{r-1} \frac{(\lambda x)^j}{j!} & x \geq 0 \end{cases}$$

$$\text{E}(X) = \frac{r}{\lambda}, \ \text{Var}(X) = \frac{r}{\lambda^2}$$

$Y_i \overset{iid}{\sim} \exp(\lambda)$

$\boxed{X = \sum_{i=1}^{r} Y_i}$ $\rightarrow$ $E(X) = E(\sum Y_i)$

Common to see $\alpha = r$ and $\beta = \cancel{\lambda}$

$\beta = \frac{1}{\lambda}$

Identifying gamma RV from word problems

- Gamma distribution with $r = 1$ is same as exponential

continuous

  - Just like Negative Binomial with $r = 1$ is same as the geometric distribution

discrete

- Similar to exponential
  - Look for time between or until events/successes
    - BUT now we are measuring time until more than 1 success
  - Look for a rate of the events over time period

Helpful R code

dgamma : $\cancel{P(X=x)}$ $f_X(x)$

Let's say we're sitting at the bus stop, measuring the time until 4 buses arrive. We know the bus comes every 10 minutes on average.

- If we want to know the probability that the 4 buses arrive in the next 50 minutes: $P(X \le 50)$

```
1 pgamma(q = 50, rate = 1/10, shape = 4)
```

```
[1] 0.7349741
```

λ r

```
1 pgamma(q = 50, scale = 10, shape = 4)
```

```
[1] 0.7349741
```

β α

- If we want to know the time, say t, where the probability of the 4 buses arriving at t or earlier is 0.35: $P(X \le t) = 0.35$

```
1 qgamma(p = 0.35, rate = 1/10, shape = 4)
```

```
[1] 29.87645
```

- If we want to know the probability that the 4 buses arrives between 30 and 50 minutes:

```
1 pgamma(q = 50, scale = 10, shape = 4) - pgamma(q = 30, scale = 10, shape = 4)
```

```
[1] 0.382206
```

- If we want to sample 20 arrival times for the 4 buses: x realized outcome

```
1 rgamma(n = 20, scale = 10, shape = 4)
```

```
 [1] 67.04894 14.44306 43.05435 12.51608 30.58018 34.94098 28.34891 26.01258
 [9] 37.60928 24.73837 32.30761 45.71904 54.58035 41.08169 18.69450 28.78482
[17] 51.66387 40.08829 56.18500 22.85659
```

Remarks

- The parameter r in a Gamma(r,λ) distribution does NOT need to be a positive integer
  - r is usually a positive integer
- When r is a positive integer, the distribution is sometimes called an Erlang(r,λ) distribution

↳ gamma w/ $r > 0$

- When r is any positive real number, we have a general gamma distribution that is usually instead parameterized by $\alpha > 0$ and $\beta > 0$, where:
  - $\alpha =$ shape parameter : same as r, the total number of events we must witness
    - In R code example: 4 buses to wait for
  - $\beta =$ scale parameter : same as λ, the rate parameter
    - In R code example: 1 bus per 10 minutes (1/10)

$\frac{1}{\lambda}$

in R : rate $= \frac{1}{\text{scale}}$

Sending money orders

Example 3

On average, someone sends a money order once per 15 minutes. What is the probability someone sends 10 money orders in less than 3 hours?

★rate★

$\lambda = \frac{1 \text{ m.o.}}{15 \text{ min}}$

$r = 10$

10 is set, but we're measuring time (time is RV)

$$\underbrace{X}_{\text{in hours}} \sim \underbrace{\text{Gamma}(r=10, \lambda = \tfrac{1}{15})}_{\text{in minutes}}$$

Two choices? NO!

- ~~scale X to be in minutes~~
- scale λ to be in hours

$$\lambda = \frac{1 \text{ m.o}}{15 \cancel{\text{min}}} \times \frac{60 \cancel{\text{min}}}{1 \text{ hr}} = \frac{4 \text{ mo}}{1 \text{ hr}}$$

$$X \sim \text{Gamma}(10, 4)$$

$$P(X \leq 3) = 0.7576$$

options:
- integrate pdf from 0 to 3
- CDF @ 3
- in R (doing above in R)

pgamma(q=3, shape=10, rate=4)

The prob that someone sends 10 money orders in less 3 hours is 0.7576.

Additional Resource

- Another helpful site with R code: https://rpubs.com/mpfoley73/459051

Normal RVs

Properties of Normal RVs

- No scenario description here because the Normal distribution is so universal
  - Central Limit Theorem (next class) makes it applicable to many types of events
- Shorthand: $X \sim \text{Normal}(\mu, \sigma^2)$

$X \sim \text{Normal}(\mu, \sigma)$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}, \text{ for } -\infty < x < \infty$$

No closed form CDF

$$\text{E}(X) = \mu$$

$$\text{Var}(X) = \sigma^2$$

$$sd(X) = \sigma$$

Helpful R code

$X \sim Norm(50, 4^2)$

Let's say we're measuring the high temperature today. The average high temperature on this day across many, many years is 50 degrees with a standard deviation of 4 degrees.

- If we want to know the probability that the high temperature is below 45 degrees:

```
pnorm(q = 45, mean = 50, sd = 4)
[1] 0.1056498
```

$P(X \leq 45) \rightarrow$ CDF

- If we want to know the temoerature, say t, where the probability of that the temperature is at t or lower is 0.35:

inverse CDF $P(X \leq t) = 0.35$

```
qnorm(p = 0.35, mean = 50, sd = 4)
[1] 48.45872
```

$\hookrightarrow 48.458$

- If we want to know the probability that the temperature is between 45 and 50 degrees:

```
pnorm(q = 50, mean = 50, sd = 4) - pnorm(q = 45, mean = 50, sd = 4)
[1] 0.3943502
```

$P(45 \leq X \leq 50)$
$= P(X \leq 50) - P(X \leq 45)$

- If we want to sample 20 days' temperature (over the years) from the distribution:

```
rnorm(n = 20, mean = 50, sd = 4)
 [1] 51.25746 45.35334 49.66657 51.78714 47.45682 53.39772 52.86730 47.12661
 [9] 49.79689 52.23243 50.95174 58.17850 51.36292 46.18399 52.27518 53.80362
[17] 51.33378 49.25484 54.50188 55.77698
```

Movie night while studying

Example 4

Children's movies run an average of 98 minutes with a standard deviation of 10 minutes. You check out a random movie from the library to entertain your kids so you can study for your test. Assume that your kids will be occupied for the entire length of the movie.

a. What is the probability that your kids will be occupied for at least the 2 hours you would like to study?

b. What is range for the bottom quartile (lowest 25%) of time they will be occupied?

let X = time kids are occupied

$X \sim \text{Normal}(\underset{\text{min}}{98}, 10^2)$

a) $\cancel{P(X \geq 2)}$ hours $\quad P(X \geq 2 \cdot 60)$

$= P(X \geq 120) = 0.013$

pnorm(q = 120, mean = 98, sd = 10, lower.tail = F)

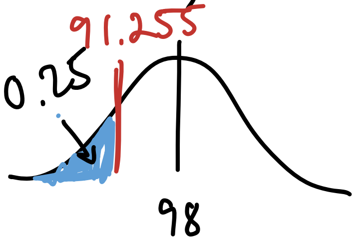
↳ $P(X > 120)$

lower.tail = T : $P(X \leq 120)$

b) $P(X \leq t) = 0.25 \quad t = ?$

= qnorm(p = .25, mean = 98, sd = 10)

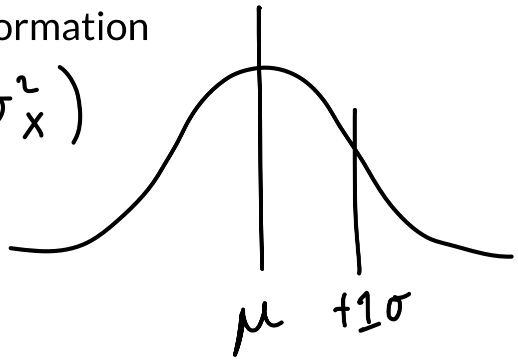
= 91.255 min



Standard Normal Distribution

$$Z \sim \text{Normal}(\mu = 0, \sigma^2 = 1)$$

- Used to be more helpful when computing was not as advanced
  - Use tables of the standard normal
  - You can convert any normal distribution to a standard normal through transformation
- $Z = \dfrac{X - \mu_X}{\sigma_X}$ } using $X \sim \text{Norm}(\mu_X, \sigma_X^2)$
  - Comes from $X = \sigma_X Z + \mu_X$
  - Since σ_X and μ_X are constants, then $E(X) = \mu_X$ and $SD(X) = \sigma_X$



$$E(X) = E(\sigma_X Z + \mu_X)$$

$$= \underbrace{E(\sigma_X Z)}_{0} + \underset{\mu_X}{E(\mu_X)}$$

Chi-squared distribution

X is always positive

- If $Z \sim \text{Normal}(0, 1)$, then $X = \boxed{Z^2}$ has a Chi-squared distribution with 1 degree of freedom
  - Shorthand: $X \sim \chi^2_{(1)}$

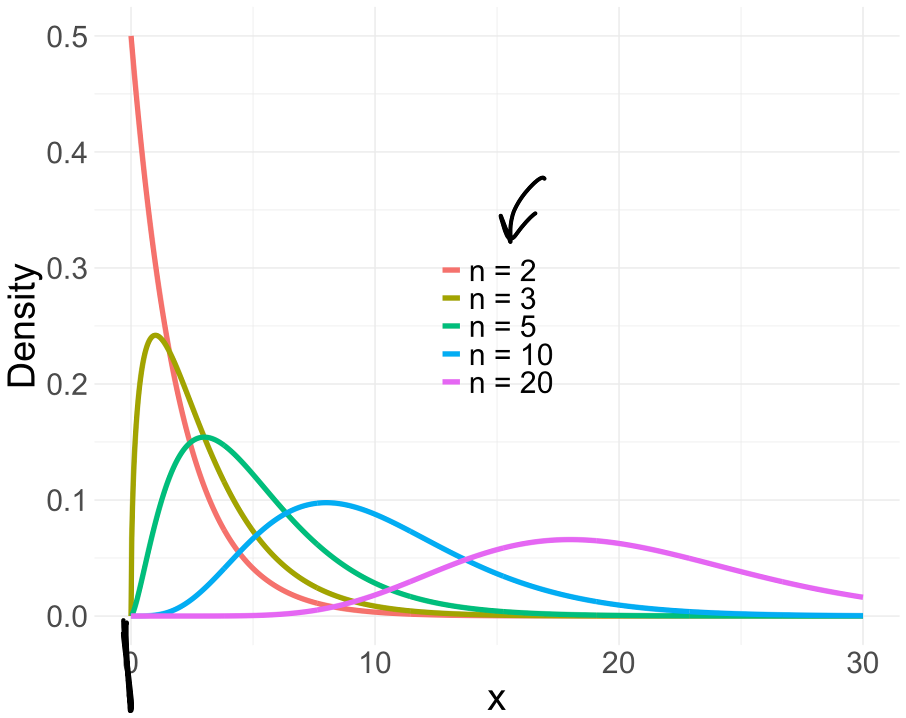
$X = \sum_{i=1}^{1} Z^2$

- If $Z_1, Z_2, \ldots, Z_n$ are independent standard normal RVs, then

$$X = \sum_{i=1}^{n} Z_i^2$$

sum of n std normals squared

has a Chi-squared distribution with n degrees of freedom
  - Shorthand: $X \sim \chi^2_{(n)}$



$\text{Normal} \xrightarrow{\Sigma Z^2} \text{Chi} \longrightarrow F$

not sq

sq

$\dfrac{\left[\chi^2_n/n\right]}{\left[\chi^2_m/m\right]}$

ratio of sq's