Welcome to CS109A

Gili Rusak

Agenda

- Expectation and Normal Distribution*
- Conditional Expectation*
- Continuous Random Variables*
- Approximations of the Binomial Random Variable*

* Relevant for Quiz 2

Exam Topics

- Expectation*
- Discrete and Continuous Random Variables*
- Normal Distribution*
- Approximations of the Binomial Random Variable*
- Joint Distributions*
- Independent Random Variables*
- Joint Random Variables*
- Conditional Expectation*
- General Inference*

* Relevant for Quiz 2

Normal Random Variable

Practice Exam Problem 1

Normal CDF

The Stanford Competitive Robotics Club is building a robot that is able to throw Marshmallow Peeps. A team's robot gets one throw.

- If it throws a distance **less than 4 meters**, the team loses and gets nothing.
- If it throws a distance between **4 and 7 meters**, the team wins $5.
- If it throws a distance between **7 and 10 meters**, the team wins $20.
- If it throws a distance **greater than 10 meters**, the team loses and gets nothing.

Stanford's current robot design throws distances that are roughly normally distributed, with $\mu = 7$ meters and $\sigma 2 = 4$ meters2 (so $\sigma = 2$ meters). (There's a very small chance the robot throws the Peep backwards: a negative distance. That still counts as a loss; you don't need to handle this case specially.) On average, how much money does the team expect to win?

Practice Exam Problem 1 Solution

Step 1: Let *W* = amount the team wins in dollars, and let *D* be the distance of the throw. *D* ~ N(7, 4).

Practice Exam Problem 1 Solution

Step 1: Let *W* = amount the team wins in dollars, and let *D* be the distance of the throw. *D* ~ N(7, 4).

Step 2: Definition of Expectation

$$E[W] = \sum_{w : p_W(w) > 0} w \cdot p_W(w)$$

Practice Exam Problem 1 Solution

Step 3: Expand E[W] as pertaining to this problem

$$
\begin{aligned}
E[W] &= \sum_{w:p_W(w)>0} w \cdot p_W(w) \\
&= 0 \cdot p_W(0) + 5 \cdot p_W(5) + 20 \cdot p_W(20) \\
&= 5 \cdot p_W(5) + 20 \cdot p_W(20)
\end{aligned}
$$

Practice Exam Problem 1 Solution

Step 4: Solve using the PDF of a Normal Distribution

$$
\begin{aligned}
&= 5 \cdot p_W(5) + 20 \cdot p_W(20) \\
&= 5 \cdot P(4 < D \leq 7) + 20 \cdot P(7 < D \leq 10) \\
&= 5 \cdot P\left(\frac{4-7}{2} < \frac{D-7}{2} \leq \frac{7-7}{2}\right) + 20 \cdot P\left(\frac{7-7}{2} < \frac{D-7}{2} \leq \frac{10-7}{2}\right)
\end{aligned}
$$

Practice Exam Problem 1 Solution

Step 5: Simplify the final expression.

$$= 5 \cdot P\left(\frac{4-7}{2} < \frac{D-7}{2} \le \frac{7-7}{2}\right) + 20 \cdot P\left(\frac{7-7}{2} < \frac{D-7}{2} \le \frac{10-7}{2}\right)$$

$$= 5 \cdot P(-1.5 < Z \le 0) + 20 \cdot P(0 < Z \le 1.5)$$

$$= 5 \cdot (\Phi(0) - \Phi(-1.5)) + 20 \cdot (\Phi(1.5) - \Phi(0))$$

$$= 5 \cdot (\Phi(0) - (1 - \Phi(1.5))) + 20 \cdot (\Phi(1.5) - \Phi(0))$$

$$= \boxed{5 \cdot (0.5 - (1 - 0.9332)) + 20 \cdot (0.9332 - 0.5)}$$

Conditional Expectation

Conditional Expectation

Example 2 Recursive Code

Consider the following code:

```
int Recurse() {
    int x = randomInt(1, 3);  // Equally likely values
    if (x == 1) return 3;
    else if (x == 2) return (5 + Recurse());
    else return (7 + Recurse());
}
```

Let Y = value returned by “Recurse". What is $E[Y]$?

Example 2 Recursive Code Solution

Step 1: Expand E[Y] out in terms of law of total expectation

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$

Example 2 Recursive Code Solution

Step 2: Calculate E[Y| X = 1], E[Y | X=2], E[Y | X=3].

Example 2 Recursive Code Solution

Step 2: Calculate E[Y| X = 1], E[Y | X=2], E[Y | X=3].

$$E[Y|X = 1] = 3$$

$$E[Y|X = 2] = E[5 + Y] = 5 + E[Y]$$

$$E[Y|X = 3] = E[7 + Y] = 7 + E[Y]$$

Example 2 Recursive Code Solution

Step 3: Combine the results from Steps 1 and 2 to solve E[Y].

$$
\begin{aligned}
E[Y] &= E[Y|X=1]P(X=1) + E[Y|X=2]P(X=2) + E[Y|X=3]P(X=3) \\
&= 3(1/3) + (5 + E[Y])(1/3) + (7 + E[Y])(1/3) \\
&= 15
\end{aligned}
$$

Other Random Variables and Bayes Theorem

Practice Exam Problem 3

Exponential RV and Bayes Theorem

Say that two different manufacturers (call them A and B) are equally likely to produce screens for laptops. The lifetimes for the screens (measured in hundreds of hours) manufactured by each company are *independently* distributed as follows:

- Manufacturer A: lifetime of screens are normally distributed: $N(20, 4)$
- Manufacturer B: lifetime of screens are exponentially distributed: Exp(1/20)

Say we bought a laptop, have used it for 18 hundred hours so far, and the screen is still working at this point in time.

At this point in time, what is the probability that manufacturer B produced the screen?

Practice Exam Problem 3 Solution

Step 1: Define your variables

Let X = lifetime of screen in our laptop.

Let event A = manufacturer A produced the screen.

Let event B = manufacturer B produced the screen.

Practice Exam Problem 3 Solution

Step 1: Define your variables

Let X = lifetime of screen in our laptop.

Let event A = manufacturer A produced the screen.

Let event B = manufacturer B produced the screen.

Step 2: Define your goal

We want to compute $P(B \mid X > 18)$

Practice Exam Problem 3 Solution

Step 2: Define your goal

We want to compute P(B | X > 18)

Step 3: Apply Bayes Theorem

$$P(B \mid X > 18) = \frac{P(X > 18 \mid B)P(B)}{P(X > 18)} = \frac{(1 - P(X \leq 18 \mid B)) \cdot 0.5}{P(X > 18)}$$

Practice Exam Problem 3 Solution

Step 4: Expand Bayes Theorem

Noting that $(X \mid B) \sim \text{Exp}(1/20)$, we have:

$$P(B \mid X > 18) = \frac{0.5\left(1 - \left(1 - e^{-\frac{18}{20}}\right)\right)}{P(X > 18)} = \frac{0.5e^{-\frac{9}{10}}}{P(X > 18)}$$

Practice Exam Problem 3 Solution

Step 5: Calculate P(X > 18)

$$
\begin{aligned}
P(X > 18) &= P(X > 18 \mid A)P(A) + P(X > 18 \mid B)P(B) \\
&= P(X > 18 \mid A)(0.5) + P(X > 18 \mid B)(0.5) \\
&= 0.5 \cdot \left(1 - P\left(\frac{X - 20}{2} \le \frac{18 - 20}{2}\right)\right) + 0.5\left[1 - \left(1 - e^{-\frac{18}{20}}\right)\right] \\
&= 0.5 \cdot (1 - P\,(Z \le -1)) + 0.5e^{-\frac{9}{10}} \\
&= 0.5 \cdot (1 - (1 - P\,(Z \le 1))) + 0.5e^{-\frac{9}{10}} \\
&= 0.5\Phi(1) + 0.5e^{-\frac{9}{10}} \\
&= 0.5 \cdot 0.8413 + 0.5e^{-\frac{9}{10}}
\end{aligned}
$$

Practice Exam Problem 3 Solution

Step 6: Combine results from Steps 4 and 5

Substituting the previously computed value for $P(X > 18)$ into the expression for $P(B \mid X > 18)$, yields the final answer:

$$P(B \mid X > 18) = \frac{0.5e^{-\frac{9}{10}}}{P(X > 18)} = \frac{e^{-\frac{9}{10}}}{0.8413 + e^{-\frac{9}{10}}}$$

Reminders

Reminders

If you see the text, "use an approximation", remember we have learned about two approximations: Normal Approximation, and Poisson Approximation.

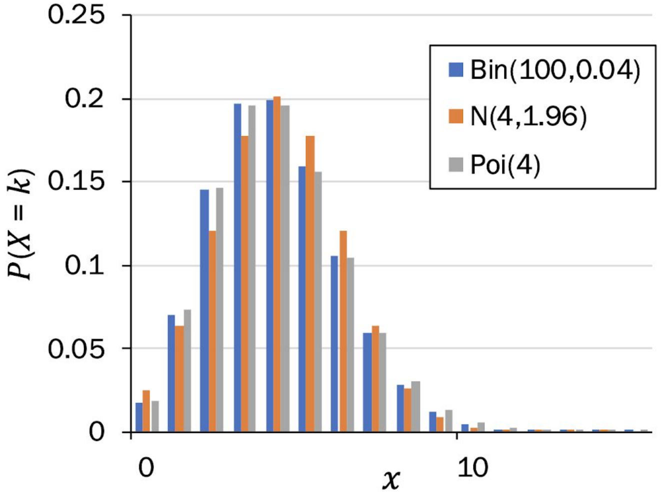
Steps:

1. Check values of n and p from your binomial distribution to determine which approximation is appropriate

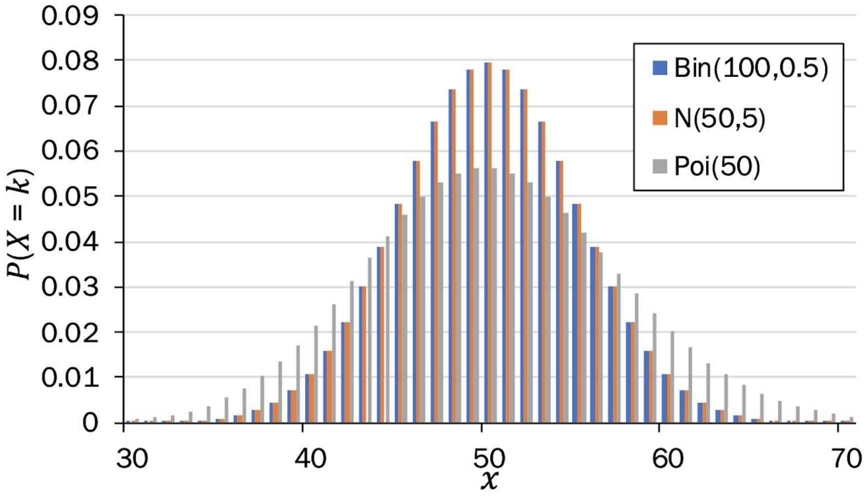
2. Apply the appropriate approximation

3. Solve the original problem

Approximations of Binomial RV



Poisson approximation
n large (> 20), p small (< 0.05)
slight dependence okay



Normal approximation
n large (> 20), p mid-ranged ($np(1-p) > 10$)
independence

1. If there is a choice, use Normal to approximate.
2. When using Normal to approximate a discrete RV, use a continuity correction.

Example Problem 4

1. Suppose the lifetimes of computer chips produced by a certain manufacturer have probability *p = 0.04* probability of lasting more than 3*10^6 hours.

What is the approximate probability that a batch of 100 chips will contain at least 6 whose lifetimes are more than 3x10^6 hours?

2. Suppose the lifetimes of computer chips produced by a certain manufacturer have probability *p = 0.67* probability of lasting more than 3*10^6 hours.

What is the approximate probability that a batch of 100 chips will contain at least 6 whose lifetimes are more than 310^6

Example Problem 4 Solution

Step 1: What is the exact binomial setup?

Step 2: Which approximation is appropriate in problem 1? Which approximation is appropriate in problem 2?

Example Problem 4 Solution

Step 1: What is the exact binomial setup?

X ~ Bin (100, 0.04)

Y ~ Bin (100, 0.67)

Step 2: Which approximation is appropriate in problem 1? Which approximation is appropriate in problem 2?

Example Problem 4 Solution

Step 1: What is the exact binomial setup?

X ~ Bin (100, 0.04)

Y ~ Bin (100, 0.67)

Step 2: Which approximation is appropriate in problem 1? Which approximation is appropriate in problem 2?

X_approx ~ Poi

Y_approx ~ Normal

Good luck on Quiz 2!