# CS109: Probability for Computer Scientists

### Lecture 11 — Inference

January 30

---

## Name to Age

Let $B$ be the year that someone was born and let $N$ be their name. From the social security dataset, I am able to give you the joint distribution $P(B = b, N = n) \approx \frac{\text{count}(b,n)}{\text{len(dataset)}}$.

(a) How would you compute $P(B = 1964 | N = \text{Michael})$ using the joint distribution?

(b) Describe how you would compute the full PMF $P(B = b | N = \text{Michael})$, over all birth years $b$, using code.

**(a)** We use the law of total probability:

$$P(X = \text{Relationship}) = \sum_y P(X = \text{Relationship}, Y = y)$$

Summing the Relationship column:

$$P(X = \text{Relationship}) = 0.08 + 0.11 + 0.10 + 0.07 + 0.09 = 0.45$$

**(b)** We again use the law of total probability:

$$P(Y = \text{Frosh}) = \sum_x P(X = x, Y = \text{Frosh})$$

Summing the Frosh row:

$$P(Y = \text{Frosh}) = 0.13 + 0.08 + 0.02 = 0.23$$

## Hidden Chambers

a. Imagine the entire 100 meter path is limestone. In that case, the rate of muons arriving per month on the detection plate is

$$100 \cdot e^{-100/40} = 8.2.$$

Assume each muon arrives independently of any other muon and at a constant rate. What is the probability that in one month you would observe 12 muons?

b. Let $X$ be your belief in the meters of limestone above the detection plate. Your prior belief is that any number of meters from 0 to 100 is equally likely:

$$X \sim \text{Uni}(0, 100).$$

After one month, your detection plate has been hit by 12 muons. What is your updated belief in $X$? You may leave your answer with integrals or sums.

a. The number of muons observed in a month follows a Poisson distribution with rate

$$\lambda = 100 \cdot e^{-100/40} = 8.2.$$

Thus,

$$P(N = 12) = \frac{\lambda^{12} e^{-\lambda}}{12!} = \frac{8.2^{12} e^{-8.2}}{12!}.$$

b. For a given value of $X = x$, the rate of muons is

$$\lambda_x = 100 \cdot e^{-x/40}.$$

The likelihood of observing 12 muons given $X = x$ is

$$P(N = 12 \mid X = x) = \frac{\lambda_x^{12} e^{-\lambda_x}}{12!}.$$

The prior density is

$$p(x) = \frac{1}{100}, \quad x \in [0, 100].$$

By Bayes' rule, the posterior density is

$$p(x \mid N = 12) \propto P(N = 12 \mid X = x)\, p(x) = \frac{(100e^{-x/40})^{12} e^{-100e^{-x/40}}}{12!} \cdot \frac{1}{100}, \quad x \in [0, 100].$$

A normalization constant can be obtained by integrating this expression over $x \in [0, 100]$.

## Stanford Eye Test

Let $A \in \{0.00, 0.01, \ldots, 0.99, 1.00\}$ be the (discretized) ability that someone can see. Our prior belief in A is given to you as a dictionary for each possible ability. In this eye test, we show a user a letter at a certain font size and we observe if they get it right or wrong. Then we update our belief in their ability to see.

We observe that a user gets the first letter we show them incorrect. We define $Y = 0$ as the event that the user gets the first letter we show them incorrect.

a. Write an expression in math for the posterior $P(A = a | Y = 0)$.

b. Is this expression going to result in a number or a dictionary?

c. Describe in code how you would solve this expression. Assume you have access to a function **calc_likelihood**. (The likelihood function is super neat - a bit outside of the scope of today's class but happy to talk about it in OH).

a. We can use Bayes Theorem to derive an expression for the updated belief in a person's ability to see, given they saw an incorrect letter:

$$
\begin{aligned}
P(A = a \mid Y = 0) &= \frac{P(Y = 0 \mid A = a)P(A = a)}{P(Y = 0)} \\
&= \frac{P(Y = 0 \mid A = a)P(A = a)}{\sum_a P(Y = 0 \mid A = a)P(A = a)}
\end{aligned}
$$

b. A dictionary!

$A$ is a discrete variable with values ranging from $\{0.00, 0.01, ..., 0.99, 1.00\}$. The expression $P(A = a|Y = 0)$ represents an entire distribution, namely the updated distribution for $A$ after conditioning on having observed the incorrect letter.

Using the expression in part a, we can iterate through every discrete value of $A$, and compute an updated probability that a person has a particular ability level $a$.

**Note:** If we were interested in computing the updated probability that a person has a particular ability level $a$ (such as 0.55), we can use Bayes Theorem to compute the updated probability. The following expression computes a number:

$$
\begin{aligned}
P(A = 0.55 \mid Y = 0) &= \frac{P(Y = 0 \mid A = 0.55)P(A = 0.55)}{P(Y = 0)} \\
&= \frac{P(Y = 0 \mid A = 0.55)P(A = 0.55)}{\sum_a P(Y = 0 \mid A = a)P(A = a)}
\end{aligned}
$$

c.
```python
def update_belief(prior):
    posterior = {}
    normalization_constant = 0.0

    for ability in prior:
        posterior[ability] = calc_likelihood(ability) * prior[ability]
        normalization_constant += posterior[ability]

    # Normalize posterior distribution
    for ability in posterior:
        posterior[ability] /= normalization_constant

    return posterior
```

## 1-D Tracking

Imagine you have a self driving car with one LiDAR sensor. A LiDAR sensor is a way to measure distance to other objects. You don't need to know how LiDAR works, just that it will give you a measure of distance. You are trying to detect how far away an object is from the self driving car.

Let $T$ be the true distance from the car to the object. Our prior belief is that $T \sim N(\mu = 1, \sigma^2 = 3)$.

Our LiDAR sensor can measure distance but it isn't perfectly accurate. It has some noise due to measurement error within the instrument. Let $X$ be the distance given by the LiDAR. We say that $X$ is equal to the true distance plus some noise: $X = t + \text{Noise}$. Let $M$ be the noise and $M \sim N(\mu = 0, \sigma^2 = 1.5)$.

a. What is the likelihood function $f(X = x|T = t)$? This is asking, if you knew a value for t, how could you express $X$. Specifically, use linearity of expectation and linearity of variance to find the parameters of this distribution.

b. We observe a LiDAR measurement of 4 meters. Write out the equation for the probability density $f(X = 4|T = t)$.

c. We want to update our belief in the true distance to the object $f(T = t|X = 4)$. Write your answer in terms of a constant K that you do not need to solve for.

a. Before deriving the PDF, a good starting point is to determine the distribution of $X$. $X$ represents a noisy LiDAR measurement. We can express $X$ as the sum of some true distance and a sample valued from the noise variable $M$ (in other words, $X = t + M$)

If we knew a value for the true distance $t$, then the distribution of $X$ is simply a linear transformation of the distribution of $M$. By adding $t$ to $M$, we shift the entire distribution of $M$ by $t$ units to right.

The linear transformation of a Normal distribution always results in a Normal distribution. So, $X$ is a Normal distribution.

We can use linearity of expectation to compute the expectation of $X$:

$$\begin{aligned} E[X] &= E[t + M] \\ &= t + E[M] \\ &= t + 0 \\ &= t \end{aligned}$$

We can use linearity of variance to compute the variance of $X$:

$$\begin{aligned} Var[X] &= Var[t + M] \\ &= 0 + Var[M] \\ &= 1.5 \end{aligned}$$

Putting it all together, $X \sim N(\mu = t, \sigma^2 = 1.5)$ and the probability density function of $X$ (which we express as $f(X = x|T = t)$) is

$$f(X = x|T = t) = \frac{1}{\sqrt{2\pi}\sqrt{1.5}} e^{-\frac{(x-t)^2}{2(1.5)}}$$

b. We input a value of 4 for $x$ in the PDF expression derived in part a:

$$f(X = 4|T = t) = \frac{1}{\sqrt{2\pi}\sqrt{1.5}} e^{-\frac{(4-t)^2}{2(1.5)}}$$

c. We can use Bayes Theorem to compute an updated belief that the true distance is $t$:

$$\begin{aligned} f(T = t|X = 4) &= \frac{f(X = 4|T = t) f(T = t)}{K} \\ &= \frac{\frac{1}{\sqrt{2\pi}\sqrt{1.5}} e^{-\frac{(4-t)^2}{2(1.5)}} \cdot \frac{1}{\sqrt{2\pi}\sqrt{3}} e^{-\frac{(t-1)^2}{2(3)}}}{K} \end{aligned}$$