Statistical Thinking for Forensic Practitioners

Part 5: Probability Models and Uncertainty (Video 1)

Dr. Kiegan Rice

Section Outline

- Measurement, variability, and uncertainty
- Reliability, repeatability and reproducibility, accuracy and precision
- Probability distributions and parameters
- Probability models for discrete variables
- Probability models for continuous variables
- Expectation, variance, covariance, and correlation of variables
- Propagation of error

Why are these concepts important?

- Section 3: Data collection
  - focused on the "big picture" of data
- Section 4: Data types and summaries
  - focused on the "basics" of data
- Section 5: Uncertainty and probability models
  - starting to get into the "nitty gritty" of using collected data
  - how concepts can be applied in forensic science

Measurement, Variability, and Uncertainty

- Any measurement process involves some degree of uncertainty and variability
- Natural variability in life and in measurement processes
  - If you measure some item multiple times, you will not get exactly the same answer
  - Example: mass of an object

ISO 1725: 7.6.1. Laboratories shall identify the contributions to measurement uncertainty. When evaluating measurement uncertainty, all contributions that are of significance, including those arising from sampling, shall be taken into account using appropriate methods of analysis.

- Measure of resulting uncertainty should be provided to the user (and reported in general)

Defining Uncertainty

- Scientists focus on physical measurements
  - Uncertainty refers to intrinsic uncertainty in a measurement
  - Example: thermometer only accurate to within 0.1 degrees
  - Significant digits / significant figures can affect this
- Statisticians focus more broadly
  - All kinds of uncertainties, including measurement, prediction uncertainty, things we are unsure about
- We can address uncertainty by using *probability*

Uncertainty and Probability

We use probability to account for uncertainty that exists in the world and in science.

Probability can be used to express uncertainty in multiple ways:

- Using a probability statement
  - "The probability of rain tomorrow is 60%."
- Using a probability distribution
  - "Weights on this scale are normally distributed with a standard deviation of 0.1 kg."
- Using summaries from probability distributions
  - "This measurement is accurate to $\pm$ 0.5 in."
- We will discuss distributions in more detail later in this section.

Variability in Measurement

- **Variability** is part of probability statements and distributions.
  - Variability is what is *observed*.
  - Uncertainty is how we *express* it.
  - Probability is how we *account* for it.
- Variability refers to variation observed in *repeated measurements*
  - repeated measurements of the same object under the same environmental conditions
  - repeated measurements of the same object under different environmental conditions
  - repeated measurements of different (but related) objects

Variability in Measurement

- Variability is calculated by measures of *spread* (Section 4)
  - Standard deviation, variance, range, IQR
- Consider glass refractive index:
  - If we measure refractive index multiple times, we might get slightly different measurements each time.
  - We might report refractive index on a glass fragment as an average measurement with a distribution.
  - $\text{RI} \sim N(1.51, 0.02)$
- This concept can be applied to all quantitative measurements.
  - Examples: mass of an object, length of an object, width of an object.

Variability in Forensic Science

- Measurement doesn't always refer to physical, quantitative measurements.
- Variability can refer to a lot of aspects of forensic science.
- Questions about uncertainty are central to thinking about forensic science.

Example:

- Suppose we think signature complexity is relevant to assessing signature evidence in forensic document examination. How can we measure complexity, and is our measurement reliable?

Some definitions:

- **Reliability**: How consistently a method measures some information.
- **Repeatability**: How consistently a method measures some quantity (or quality) in the same environmental conditions (same operator/person, same tool, etc.).
- **Reproducibility**: How consistently a method measures some quantity (or quality) across environmental conditions (different operators/people, different tools, etc.).

Repeatability and reproducibility together make up a method's reliability.

Let us consider the example of handwriting complexity:

- Suppose we want to measure how complex a signature is. "Complexity" is difficult to define numerically/quantitatively.

Consider the following study:

- Five forensic document examiners (FDEs) rated 123 signatures in terms of difficulty on a 5-point scale: (1) easy, (2) fairly easy, (3) medium, (4) difficult, (5) very difficult.

ID	FDE1	FDE2	FDE3	FDE4	FDE5
001	4	4	5	3	4
002	4	5	5	4	5
003	3	4	4	4	3
004	4	4	5	4	4
005	2	2	2	3	3
...	...	...	...	...	...

These data can be used to assess *reproducibility* of complexity ratings by FDEs.

Suppose in addition we want to measure the *repeatability* for the complexity ratings.

- Using the same signatures used in the previous study, have some examiners rate a subset of signatures multiple times.
- Each examiner repeats ratings for those signatures.
- How variable are the ratings for an examiner?

In this particular example, we can also think of these as **inter-rater reliability** (across FDEs) and **intra-rater reliability** (for each FDE).

- What makes a process "reliable"? Standards depend on the field and the way the study is designed.
- One commonly used threshold: if over 10% of total variation in a study can be attributed to differences in operators or repetition, the measurement process is not reliable.

The data collected on signature complexity helps us think about several concepts:

- Experimentation and data collection are an important part of thinking about uncertainty.
- Repetition in measurements is critical to show that the procedure is repeatable and reproducible.
- How experiments are designed and carried out affects what we can learn from them.
  - Ratings from multiple examiners = reproducibility of rating.
  - Multiple ratings from same examiner = repeatability of rating.

Other considerations:

- Overall reliability of a method considers repeatability and reproducibility together.
- Often, the ratio of repeatability and reproducibility is also of interest.

Practical implications of studying reliability:

- Measurement:
  - How reliable are measurement tools when used by the same operator?
  - How reliable is the application of those tools by different operators?
- Ratings and Evaluations:
  - How consistent are evaluations of the same evidence or object by a single examiner?
  - How consistent are evaluations of the same evidence or object by multiple examiners?

These help us address questions about reliability of forensic evaluation in general.

To recap...

- Uncertainty is inherent in measurement and science.
- Repeated measurements can be used to assess repeatability and reproducibility.
- Different study designs allow us to answer different types of questions about reliability.

Statistical Thinking for Forensic Practitioners

Part 5: Probability Models and Uncertainty (Video 2)

Dr. Kiegan Rice

Center for Statistics and Applications in Forensic Evidence

We use probability to incorporate the uncertainty we know exists in measurement into assessments and conclusions. To do this, we use **probability distributions**. A probability distribution is made up of:

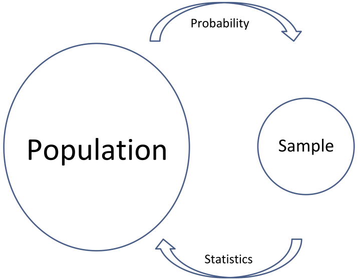
- Possible values.
- How likely each value is to occur.

Consider a measurement of interest, such as glass refractive index or "complexity" of a signature.

- We want to know the *true value* of the measurement of interest. We call this the **parameter** of interest, θ.
- θ is the "true" complexity of a signature.
- We sample data (e.g., $x_1, x_2, x_3, \ldots$) to estimate what that value is. Here samples are ratings by forensic examiners.

Probability Distributions and Parameters

We think about probability distributions as enumerating the possible values that θ could take on, informed by our observations of the "real world":



This is important for thinking about how likely evidence is under a certain hypothesis!

Probability distributions all have the following properties:

- For any possible value X, $P(X) \in [0, 1]$.
  - $P(X) \geq 0$
- If Ω is the set of all possible values (this is S in Section 2),

$$\sum_{X \in \Omega} P(X) = 1$$

Thinking about Ω:

- Coin flip: $\Omega = \{H, T\}$
  - I flip. I observe $x = H$.
- Die roll: $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - I roll. I observe $x = 2$.
- Mass of an object in grams: $\Omega = (0, \infty)$
  - I weigh a bullet. I observe $x = 7.03g$.

We will define some specific probability distributions, in two categories:

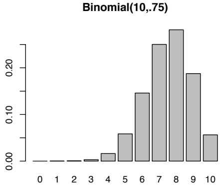
- **Discrete distributions:** Possible values $X \in \Omega$ take on discrete values. We can typically enumerate all possible values of X.
  - Examples: coin flip, dice roll
- **Continuous distributions:** Possible values $X \in \Omega$ can take on any value within a range of possible values. We typically cannot enumerate all possible values of X.
  - Example: refractive index, mass of object
  - *Important note:* Probability of any single value is zero. $P(X = 7.03) = 0$. However, we can calculate: $P(7.02 < X \leq 7.03) = 0.05$.

Binomial Distribution

- *Distribution type:* Discrete
- *Data structure it captures:* Number of successes in n trials.
- *Example:* If I flip a coin $n = 10$ times, X might be how many times I get heads.
- *Forensic example:* If I test $n = 10$ bags of contraband, X might be how many bags contain drugs.
- *Definition:* $X \sim Binom(n, p)$. For some value k,

$$P(X = k) = \binom{n}{k}(p)^k(1-p)^{n-k};$$

$$p = \text{prob. of success}, n = \text{number of trials}$$



Binomial Distribution: Example

Let's consider an example and calculate the probability of interest:

- Suppose ten bags of contraband are confiscated, and need to be tested for the presence of drugs.
- Suppose the probability of a single bag containing drugs is $p = 0.75$.
- What is the probability of 6 bags testing positive for drugs?
  - Recall:
$$P(X = k) = \binom{n}{k}(p)^k(1-p)^{n-k};$$
  - $p = 0.75,\ n = 10,\ k = 6$:

$$P(X = 6) = \binom{10}{6}(0.75)^6(1-0.75)^{10-6}$$

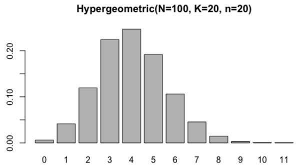
$$= \frac{10!}{6!(10-6)!}(0.75)^6(1-0.75)^{10-6} = 0.146$$

Hypergeometric Distribution

- *Distribution type:* Discrete
- *Data structure it captures:* Number of successes in n trials without replacement out of a population of N objects, of which K objects have the feature of interest.
- *Example:* A jar (or bag or urn) with two colors of marbles, N total. If I draw n marbles, X might be the number of red marbles drawn.
- *Forensic example:* A shipment of contraband contains N teddy bears. Suppose K bears contain illegal drugs. If I draw n teddy bears and test each for the presence of drugs, X might be the number that test positive.
- *Definition:* $X \sim \text{Hypergeometric}(N, K, n)$. For some value k,

$$P(X = k) = \frac{\binom{K}{k}\binom{N-k}{n-k}}{\binom{N}{n}};$$

$N =$ total objects in population, $K = \#$ of successes in the population, $n = \#$ of draws

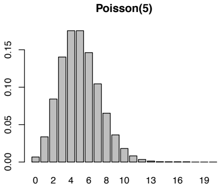


Poisson Distribution

- *Distribution type:* Discrete
- *Data structure it captures:* Count of number of events occurring in some interval of space or time.
- *Example:* How many patients arrive at an emergency room in a given hour.
- *Forensic example:* If I compare striae on two land engraved areas (LEAs), X might be the number of consecutively matching striae.
- *Definition:* $X \sim Pois(\lambda)$. For some value k,

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!};$$

$$\lambda = \text{mean number of events in an interval}$$

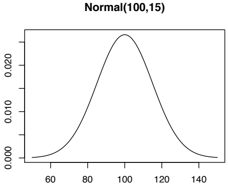


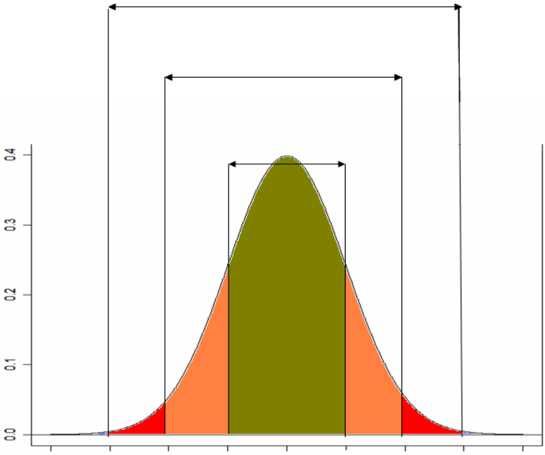
Normal Distribution

- *Distribution type:* Continuous
- *Data structure it captures:* Bell-shaped curve.
- *Examples:* Heights, weights, blood pressure.
- *Forensic example:* If I measure packages of drugs found on a suspect, X might be the weight of a package.
- *Definition:* $X \sim N(\mu, \sigma^2)$. For some value x,

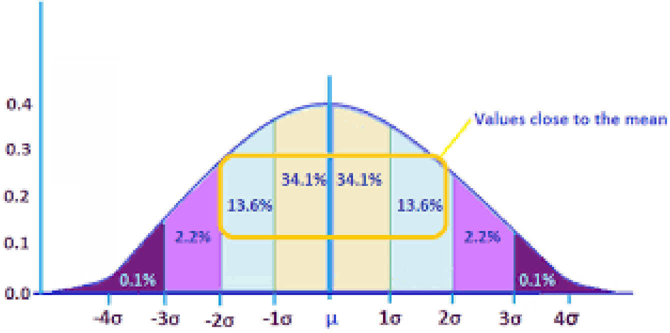
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\mu = \text{mean of dist.}, \sigma^2 = \text{variance of dist.}$$




0.4
0.3
0.2
0.1
0.0

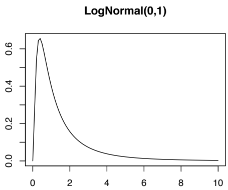
Normal Distribution



Log-Normal Distribution

- *Distribution type:* Continuous
- *Data structure it captures:* Logarithm of observations follow a normal distribution.
- *Forensic example:* If I measure concentration of a chemical in glass fragment, X might be the concentration.
- *Definition:* $X \sim \text{Lognormal}(\mu, \sigma^2)$. For some value x,

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$$



To recap...

- Probability distributions help us capture uncertainty in measurement into analysis.
- Probability distributions are defined using parameters.
- Commonly used probability distributions and their parameters.

Up next...

- More on parameters
- Expectation, variance, covariance, and correlation
- Accuracy and precision

Statistical Thinking for Forensic Practitioners

Part 5: Probability Models and Uncertainty (Video 3)

Dr. Kiegan Rice

More on Parameters

Probability distributions are useful to capture uncertainty and variability in measurement and evaluation using formal probability statements.

Often, we are specifically interested in certain values associated with the distribution, because they capture quantities that are useful.

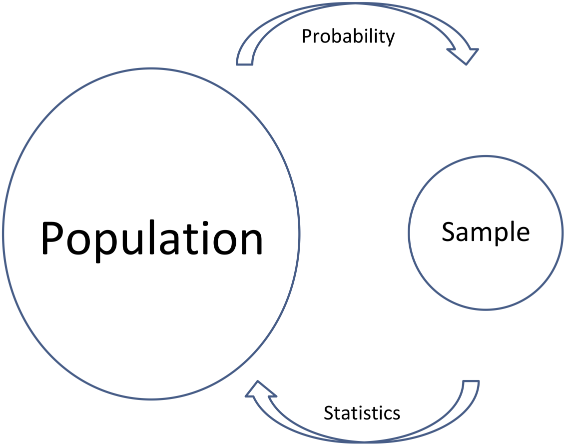
- In Section 4, we discussed measures of *center* and *spread*.
- We will talk about these measures as they relate to probability distributions: **expectation** and **variance**.

More on Parameters

Some notation:

- When we take a sample of size n $(x_1, x_2, \ldots, x_n)$, we can use the sampled values to estimate the value of parameters.
- If our data follow a normal distribution, we might be interested in estimating μ (population mean) and σ^2 (population variance).
  - *Note:* Often we are more interested in σ (population standard deviation).
- We can estimate μ by calculating the **sample mean**, $\bar{x}$.
- Depending on the distribution, we might use the data in different ways to estimate different parameters. (This will be covered more in Section 6).
- However, we can think about some theoretical values related to our distribution that are useful for estimating parameters.

More on Parameters



Expectation

Expectation

The **expected value** of a random variable X is denoted as $E(X)$.

- E(X) is the value we would "expect" if drawing at random from the probability distribution of X.
- E(X) is defined as:

$$\text{Discrete: } \sum_{x \in \Omega} x \; p(x)$$

$$\text{Continuous: } \int_{x \in \Omega} x \; f(x) \; dx$$

- We can think of expected value as the weighted average of all possible values of x, weighted by their respective probabilities.
- Expected value is a measure of *center*.
- For a normal random variable, $E(X) = \mu$.

Sometimes the expected value corresponds directly to a distribution parameter, and sometimes it does not.

Expected values for the distributions previously presented are:

Distribution	**E(X)**
Binomial	$n * p$
Hypergeometric	$n * \frac{K}{N}$
Poisson	λ
Normal	μ
Log-Normal	$e^{\mu + \frac{1}{2}\sigma^2}$

Example: Hypergeometric with $K = 20$, $N = 100$, and a sample of $n = 20$.

$$E(X) = n\frac{K}{N} = 20 * \frac{20}{100} = 4.$$

Variance

The **variance** of a random variable X is denoted as $Var(X)$.

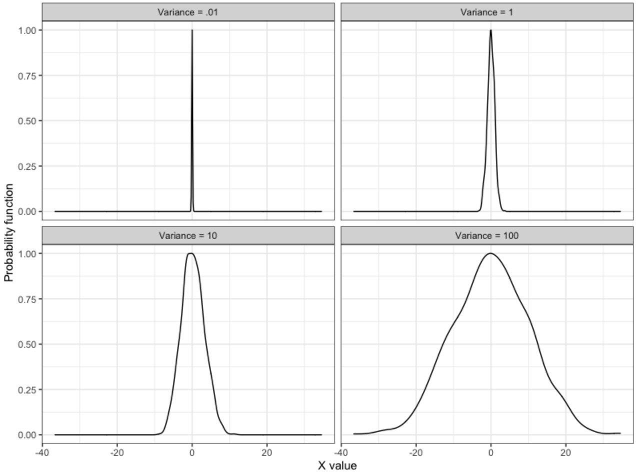
- Var(X) is a measure of *spread* for the probability distribution of X.
- Var(X) is defined as:

$$E[(X - E[X])^2]$$

- We can think of variance as the average squared deviation from the expected value of a probability distribution.

Variance

Consider the following examples of normal distributions with different variances:



Variance

The variance also sometimes directly corresponds to parameters in the distribution, but not always:

Distribution	**E(X)**	**Var(X)**
Binomial	$n * p$	$n * p * (1 - p)$
Hypergeometric	$n * \frac{K}{N}$	$n * \frac{K}{N} \frac{(N-K)}{N} \frac{N-n}{N-1}$
Poisson	λ	λ
Normal	μ	σ^2
Log-Normal	$e^{\mu + \frac{1}{2}\sigma^2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$

Consider the binomial example. If I test $n = 10$ bags of contraband, and the probability of each bag containing drugs is $p = 0.75$,

$$E(X) = n * p = 10 * 0.75 = 7.5$$

$$Var(X) = n * p * (1 - p) = 10 * 0.75 * 0.25 = 1.875$$

We expect 7.5 bags to contain drugs, with a variance of 1.875 bags.

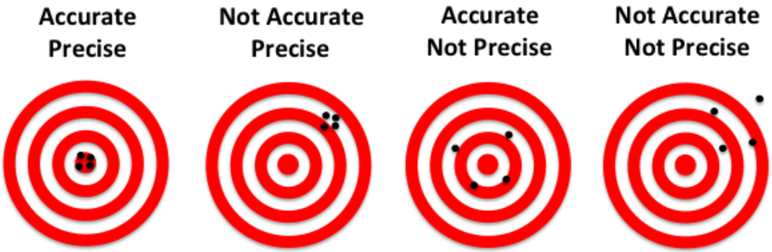
Accuracy and Precision

When we collect data (e.g. x_1, x_2, x_3), to estimate the value of a parameter, we also want to consider whether the estimates are accurate and precise.

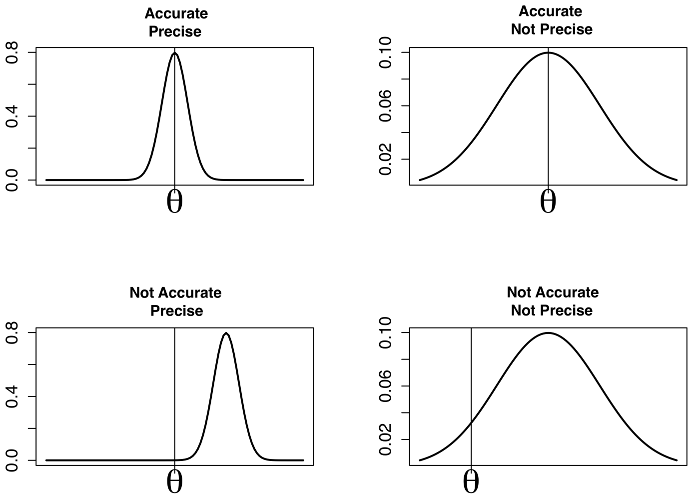
We can evaluate the estimator by considering certain properties:

- **Accuracy:** whether observed values, on average, are close to the true parameter value.
  - If observed values, on average, are not similar to the true parameter value, there is *bias*.
- **Precision:** how closely distributed values are to one another.
  - If observed values have *high precision*, they have *low variance*.

Accuracy and Precision



Accuracy and Precision



Accuracy and Precision in Practice

Q: How can we make measurements more precise?

- Take a larger sample.
- Take multiple samples and investigate the mean across samples (sampling distributions, in Section 6).

Q: When might we get an inaccurate (or **biased**) estimate?

- Data collected is from a biased sample.
  - I am interested in the distribution of signature complexity in the US population. I have five forensic document examiners rate the complexity of all signature forgery cases investigated in the last year.
  - This sample is not representative!
- Trying to capture data structure using the wrong probability distribution

Covariance

Suppose we are interested in the relationship between two random variables X and Y.

We can look at the **covariance**, or a measure of how two random variables change together.

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))]$$

Covariance measures the extent to which the deviation of a variable from its mean matches the deviation of the other variable from its mean. (The strength of the relationship).

$Cov(X, Y) \in (-\infty, \infty)$

Example: Concentration of two elements in glass fragments.

Correlation

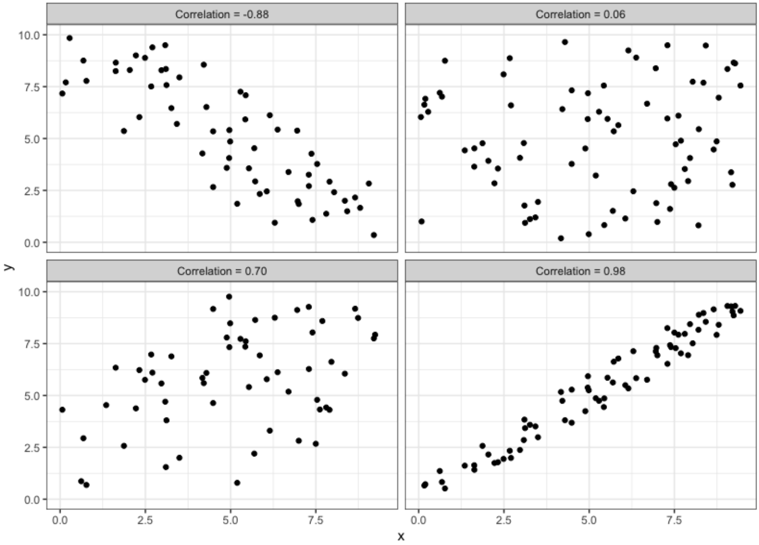
Correlation also measures the strength of the relationship between two variables, X and Y. However, it accounts for the scale and variability of each variable.

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

$$Corr(X, Y) \in [-1, 1]$$

- Positive correlation $\in (0, 1]$ indicates a positive linear relationship between X and Y; as X increases, Y increases, or as X decreases, Y decreases.
- Negative correlation $\in [-1, 0)$ indicates a negative linear relationship: as X increases, Y decreases, or as X decreases, Y increases.
- Correlation of 0 indicates no linear relationship between X and Y.

Correlation



Some additional properties:

- If X and Y are independent, $Cov(X, Y) = 0$ and $Corr(X, Y) = 0$.
- $Var(aX) = a^2 Var(X)$ for some value a.
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$ for some values a, b.

Propagation of Error

If we are interested in estimating a function of two random variables we know, we have to think carefully about expressing the variability surrounding that estimate.

- Suppose we are interested in measuring the percentage of antimony in a bullet.
- The measured antimony (X) is 2g with a standard deviation of .5g.
- The bullet has a measured weight (Y) of 64g, with a standard deviation of 1g.
- If we estimate the proportion of antimony in the bullet, $\frac{X}{Y}$, how variable is that estimate?

$$Var\left(\frac{X}{Y}\right) = \left(\frac{E(X)}{E(Y)}\right)^2 \left(\frac{Var(X)}{E(X)} + \frac{Var(Y)}{E(Y)}\right) = \left(\frac{2}{64}\right)^2 \left(\frac{.5^2}{2^2} + \frac{1^2}{64^2}\right)$$

Let's think about why this is:

- We are uncertain about the value of X (measured antimony)
- We are uncertain about the value of Y (measured weight of bullet)
- To estimate the variance of a function of the two, we have to account for how uncertain we are of *both quantities*.

Deriving the formulas such as the one used on the last slide requires some calculus, so we will skip that for now! Here are some of the common functions and variance formulas:

Function	**Approximate variance**
$\frac{1}{X}$	$\frac{\sigma_X^2}{\mu_X^4}$
$\ln X$	$\frac{\sigma_X^2}{\mu_X^2}$
$^{*}e^X$	$e^{2\mu_X}\sigma_X^2$

* can produce a highly skewed distribution so the approximation may be inaccurate. It is particularly true for X having a large variance.

Function	Approximate variance if X and Y are independent	Term to be added if X and Y are correlated
$\frac{X}{Y}$	$\left(\frac{\mu_X}{\mu_Y}\right)^2\left(\frac{\sigma_X^2}{\mu_X^2}+\frac{\sigma_Y^2}{\mu_Y^2}\right)$	$\left(\frac{\mu_X}{\mu_Y}\right)^2\left(-2\frac{\sigma_{XY}}{\mu_X\mu_Y}\right)$
$\frac{X}{X+Y}$	$\left(\frac{1}{\mu_X+\mu_Y}\right)^4\left(\mu_Y^2\sigma_X^2+\mu_X^2\sigma_Y^2\right)$	$\left(\frac{1}{\mu_X+\mu_Y}\right)^4\left(-2\mu_X\mu_Y\sigma_{XY}\right)$

Putting it all together

- All measurement processes and scientific processes inherently contain variability and uncertainty.
- We can use repeated measurements to gauge the repeatability and reproducibility of a measurement process.
- Probability models are used to incorporate uncertainty into analyses and conclusions.
- We gather samples to estimate parameters (more on this later).
- Expectation, variance, covariance, and correlation can tell us a lot about a random variable and its relationship to other random variables.

Conclusion

In this section, we discussed:

- Measurement, variability, and uncertainty
- Reliability in forensic science
- Probability distributions and parameters
- Discrete and continuous variables
- Expectation and variance
- Covariance and correlation
- Propagation of error (briefly)