2 COURSE

Probability Laws

Introduction. Probability laws are essential in agricultural sciences to quantify and predict variability in biological and crop processes. They allow agronomists to analyze data, make informed decisions, and model uncertainty in areas such as genetics, yield estimation, disease occurrence, and environmental factors affecting plant production.

1 General Concepts

Definition

A probability is a function $P : \Omega \to [0, 1]$ such that for every event $A \in \Omega$:

1. $P(A) \geq 0$,
2. $P(\Omega) = 1$,
3. If $A \cap B = \varnothing$, then $P(A \cup B) = P(A) + P(B)$.

Elementary properties

1. $P(\overline{A}) = 1 - P(A)$.
2. $P(\varnothing) = 0$.
3. If $A \subset B$ then $P(A) \leq P(B)$.
4. $0 \leq P(A) \leq 1$.
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Example

Question: In a wheat field, the probability that a seed germinates is $P(A) = 0.9$. What is the probability that it does not germinate?

Model Answer (detailed):

1. The complement event $\overline{A}$ denotes "seed does not germinate".
2. By the complement rule: $P(\overline{A}) = 1 - P(A)$.
3. Substitute $P(A) = 0.9$: $P(\overline{A}) = 1 - 0.9 = 0.1$.
4. Interpretation: There is a **10%** chance that a randomly selected seed will fail to germinate.

Conditional Probability

For events A and B with $P(B) \neq 0$,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Example

Scenario: $A =$ "seed germinates", $B =$ "seed treated with growth regulator". Suppose $P(A \mid B) = 0.9$ for treated seeds and $P(A \mid B^c) = 0.7$ for untreated seeds.

Question: What does $P(A \mid B) = 0.9$ mean and how is it used?

Model Answer (detailed):

1. $P(A \mid B) = 0.9$ means that **given** the seed was treated, the probability of germination is 90%.
2. If we want the joint probability $P(A \cap B)$ and we know $P(B)$ (proportion of seeds treated), use $P(A \cap B) = P(B)\, P(A \mid B)$.
3. Example numeric use: if 40% of seeds are treated, $P(B) = 0.4$, then $P(A \cap B) = 0.4 \times 0.9 = 0.36$ (36% of all seeds are treated-and-germinated).
4. Interpretation: conditional probabilities allow us to separate treatment effects from population-level probabilities.

Law of Total Probability

If $B_1, \ldots, B_n$ form a partition of Ω, then for any event A,

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A \mid B_i).$$

Example

Question: A pest detection event A depends on humidity levels B_1 (low), B_2 (medium), B_3 (high). Given $P(B_1) = 0.3, P(B_2) = 0.4, P(B_3) = 0.3$ and $P(A \mid B_1) = 0.1, P(A \mid B_2) = 0.3, P(A \mid B_3) = 0.6$, compute $P(A)$.

Model Answer (detailed):

$$\begin{aligned} P(A) &= \sum_{i=1}^{3} P(B_i) P(A \mid B_i) \\ &= 0.3 \cdot 0.1 + 0.4 \cdot 0.3 + 0.3 \cdot 0.6 \\ &= 0.03 + 0.12 + 0.18 \\ &= 0.33. \end{aligned}$$

Interpretation: Overall probability of pest detection is **33%** across the field, accounting for humidity distribution.

Bayes' Theorem

For events A, B with $P(B) > 0$,

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)}.$$

If $\{A_i\}$ is a partition,

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_i P(A_i)P(B \mid A_i)}.$$

Example

Question: Let $A =$ "tomato plant infected (prevalence $P(A) = 0.05$)"; $B =$ "yellow leaf spots observed". Suppose $P(B \mid A) = 0.9$ and marginal $P(B) = 0.1$. Compute $P(A \mid B)$.

Model Answer (detailed):

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)} = \frac{0.05 \times 0.9}{0.1} = \frac{0.045}{0.1} = 0.45.$$

Interpretation: Given yellow spots, probability the plant is actually infected is **45%**. This shows how a rare disease (5%) with a sensitive but not perfectly specific symptom leads to moderate posterior probability.

Independence

Events A and B are independent if $P(A \cap B) = P(A)P(B)$. Equivalently, $P(A \mid B) = P(A)$ when $P(B) > 0$.

Example

Question: Let $A =$ "soil rich in nitrogen", $B =$ "daily sunlight > 6 hours". If $P(A) = 0.7$ and $P(B) = 0.8$ and they are independent, find $P(A \cap B)$.

Model Answer (detailed):

$$P(A \cap B) = P(A)P(B) = 0.7 \times 0.8 = 0.56.$$

Interpretation: 56% of locations (or days) will have both high nitrogen and sufficient sunlight simultaneously.

2 Discrete Probability Laws

2.2.1 Bernoulli Law

Definition

A Bernoulli random variable X with parameter p takes values:

$$P(X=1)=p, \qquad P(X=0)=1-p.$$

Expectation $E(X)=p$, variance $\text{Var}(X)=p(1-p)$.

Example

Question: Each seed has $p=0.9$ probability to germinate. What are $P(X=1)$, $E(X)$ and $\text{Var}(X)$?

Model Answer (detailed):

- $P(X=1)=0.9,\ P(X=0)=0.1$ by definition.
- Expected germination rate $E(X)=p=0.9$ (i.e. 90%).
- Variance $\text{Var}(X)=p(1-p)=0.9\times 0.1=0.09$.
- Standard deviation $\sigma=\sqrt{0.09}=0.3$.

2.2.2 Binomial Law

Definition

If X counts successes in n independent Bernoulli(p) trials, then

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}, \quad k=0,\dots,n,$$

with $E(X)=np$ and $\text{Var}(X)=np(1-p)$.

Example

Question: Out of $n=10$ tomato seeds with $p=0.8$, compute $P(X=8)$ (exactly 8 germinate). Provide a step-by-step calculation and the numerical approximation.

Model Answer (detailed):

1. Use binomial formula:
$$P(X=8)=\binom{10}{8}0.8^8(0.2)^2.$$

2. Compute the combinatorial factor:
$$\binom{10}{8}=\frac{10!}{8!2!}=\frac{10\times 9}{2}=45.$$

3. Powers:
$$0.8^8 = (0.8^4)^2 = 0.4096^2 = 0.16777216, \quad 0.2^2 = 0.04.$$
4. Multiply:
$$P(X = 8) = 45 \times 0.16777216 \times 0.04.$$
First $0.16777216 \times 0.04 = 0.0067108864$. Then $45 \times 0.0067108864 \approx 0.3019899$.
5. Round to a sensible precision: $P(X = 8) \approx 0.302$.
6. Check expectation: $E(X) = np = 10 \times 0.8 = 8$, so observing exactly 8 is a likely outcome.

3 Continuous Probability Laws

2.3.1 Normal (Gaussian) Law

Definition

A random variable X is normally distributed $X \sim \mathcal{N}(\mu, \sigma^2)$ if its density is
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

2.3.2 Standard Normal Law

Definition

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$.

Example

Question: Apple weights follow $X \sim \mathcal{N}(200, 25^2)$ (units: grams). Compute $P(175 \leq X \leq 225)$ with detailed steps.

Model Answer (detailed):

1. Standardize to $Z \sim \mathcal{N}(0, 1)$:
$$Z = \frac{X - \mu}{\sigma}.$$
2. Transform the interval endpoints:
$$z_{\text{low}} = \frac{175 - 200}{25} = -1, \quad z_{\text{high}} = \frac{225 - 200}{25} = 1.$$
3. Then
$$P(175 \leq X \leq 225) = P(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1).$$
4. Use symmetry: $\Phi(-1) = 1 - \Phi(1)$, so result $= 2\Phi(1) - 1$.
5. Numerical value: $\Phi(1) \approx 0.8413447$, hence
$$P \approx 2 \times 0.8413447 - 1 = 0.6826894 \approx 0.6827.$$

6. Interpretation: About **68.27%** of apples weigh between 175 g and 225 g (one standard deviation around mean).

Example

Question: Maize plant height: $\mu = 180$ cm, $\sigma = 15$ cm. What proportion of plants exceed 200 cm?

Model Answer (detailed):

1. Compute Z for 200 cm:

$$Z = \frac{200 - 180}{15} = \frac{20}{15} = 1.\overline{3} \approx 1.3333.$$

2. Use standard normal table or calculator:

$$P(X > 200) = P(Z > 1.3333) = 1 - \Phi(1.3333).$$

3. From tables or calculator, $\Phi(1.3333) \approx 0.9082$, so

$$P(X > 200) \approx 1 - 0.9082 = 0.0918.$$

4. Interpretation: Approximately **9.18%** of plants exceed 200 cm.

4 Laws Derived from the Normal Law

2.4.1 Chi-square Law

Definition

If $X_1, \ldots, X_n$ are independent $\mathcal{N}(0,1)$ variables, then

$$Y = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2,$$

a chi-square distribution with n degrees of freedom.

2.4.2 Student's *t* Law

Definition

Let $X \sim \mathcal{N}(0,1)$ and $Y \sim \chi_n^2$ independent. Then

$$T = \frac{X}{\sqrt{Y/n}} \sim t_n,$$

the Student's t distribution with n degrees of freedom.