APPENDIX A

Brief Outline Reviewing Probability with Application to Bayes' Theorem and Habitat Models

Brief Outline Reviewing Probability with Application to Bayes' Theorem and Habitat Models

I. Notation

A. What is the probability that something is true? This *something* is an **event.**
   1. $P(A)$ = "probability of event A occurring"
   2. $P(\sim A)$ = "probability of event A **not** occurring" (the **complement** of event A)

B. The sum or **UNION** of 2 events includes outcomes in either or both events
   1. $P(A) + P(\sim A) = P(A \cup \sim A) = U$, the **universe** (absolutely certain)
   2. $\sim U = \emptyset$, since $P(U) = 1$, $P(\emptyset) = 0$. $\emptyset$ is an **impossible** event

C. Basic properties of probability
   1. If A is any event, then $P(A) \geq 0$.
   2. If U is the largest event, then $P(U) = 1$.
   3. If events A and B have no outcomes in common, the $P(A \cup B) = P(A) + P(B)$

D. The implications of these properties include
   1. $P(A) + P(\sim A) = U = 1$
   2. $P(A) = 1 - P(\sim A)$

II. Basic Properties Applied To Habitat Models

A. Classify a **parcel of land** into one of two or more mutually exclusive **habitat** categories:
SUITABLE HABITAT *or* **MARGINAL HABITAT**
   1. If $P(S)$ is the probability that the **parcel** is **suitable habitat**, then
   2. $P(\sim S) = P(M)$, the probability that it is **marginal habitat**
   3. $P(S) = 1 - P(\sim S) = 1 - P(M)$
   4. $P(S \cup M) = P(S) + P(M) = 1$ (we are certain the **parcel** is some kind of habitat)

B. Assigning probabilities
   1. Probability as Long-Run Frequency
      The **Long-Run Frequency** of an event is the proportion of the time it occurs in a long sequence of trials, or observations (where the true probability is the **limit** of the relative frequency as the number of observations increase indefinitely)
   2. Probability as Degree of Belief
      A probability based on **Degrees of Belief** is a subjective assessement concerning whether the event in question will occur (or has occurred). (where probability is a measure of uncertainty of what is known about a population that can only be sampled, that is, only a subset of the population can be observed)

III. Degrees of Belief in Habitat Models

A. Example
   1. Biologists have made frequent observations of mule deer within an area during several winters (defined as *suitable* **winter habitat)**
   2. Studies of sage grouse show that they most often nest within 2 miles of a lek (defined as *suitable* **nesting habitat)**

B. Possible outcomes for any particular parcel of land:
   1. *Suitable* as **winter habitat** (or *suitable* **nesting habitat)**
   2. *Marginal* as **winter habitat** (or *marginal* **nesting habitat)**

C. With no further knowledge about where the parcel is located, both possible outcomes are **equally likely**, that is
   1. $P(S) = P(M) = 0.5$

D. But if you know that the parcel is within *suitable* **winter habitat** (or *suitable* **nesting habitat**), then your **Degree of Belief** is
   1. $P(S) > P(M)$

E. But, not every parcel within the habitat area is suitable, with absolute certainty. So
   1. $0.5 < P(S) < 1$
   2. meaning that we are willing to say that every parcel of land within an area of **winter range** (or **nesting habitat**) has the same probability of being suitable, somewhere between 0.51 and 0.99.
   3. In the absence of additional information, we might believe that
      a. $P(S) = 0.75$, so that $P(M) = 0.25$
      b. or $P(S) = 0.60$, $P(M) = 0.40$, or $P(S) = 0.80$, $P(M) = 0.20$

IV. A basic problem of research, knowledge, and statistics: We use information in a sample to make inferences about a population. But, new information about the population may become available, meaning that probabilities can be updated.

A. Effects of new information
   1. Initially, I am willing to say that every parcel of land within an area of **winter range** (or **nesting habitat**) has the same probability of being *suitable*, say $P(S) = 0.75$, and $P(M) = 0.25$
   2. But then, I evaluate the vegetation, V, on a specific parcel only to find that there is no vegetation (its a parking lot or well pad = $\sim V$) and I know that $\sim V$ has limited value as **winter** (or **nesting**) **habitat**
   3. Now, $P(S)$ occurring **given** that $\sim V$ has occurred is < 0.75

B. This is a **conditional probability**, written as $P(S \mid \sim V)$
   1. But another parcel has vegetation, V (maybe dense sagebrush) and now, $P(S)$ occurring **given** that V has occurred is now > 0.75
   2. Now I can say that this **conditional probability**, $P(S \mid V) > P(S \mid \sim V)$

C. **Conditional probabilities** are a key component of **Bayes' Theorem** and are related to **Joint Probabilities**

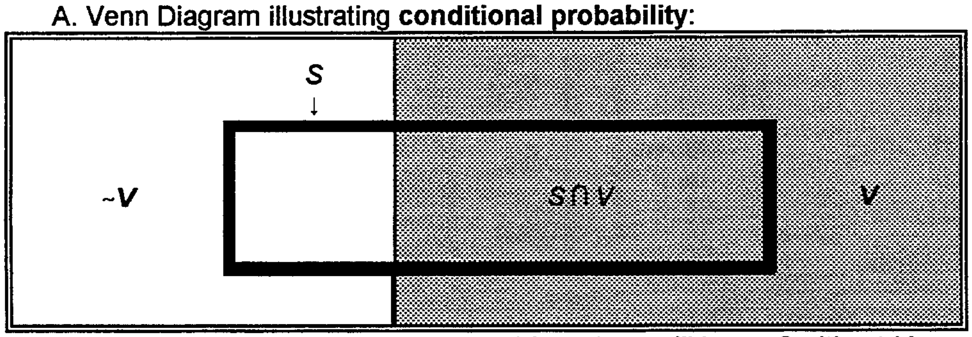
V. **Joint Probabilities** of several events occurring simultaneously

A. The simultaneous occurrence of events A and B is their **intersection**:
   1. **Intersection** of A and $B = A \cap B$ contains outcomes that are in both A and B
   2. $P(A \cap B) = P(A) \times P(B)$

B. If events A and B are **mutually exclusive** or **disjoint**
   1. $P(A \cap B) = \varnothing$

VI. **Conditional Probability of S given V**

A. Venn Diagram illustrating **conditional probability**:



   1. first requirement: that V occurs (since there will be no S without V).
   2. second, S and V intersect (a **Joint Probability**) : $P(S \cap V) = P(V) \times P(S \mid V)$
   3. which can be rewritten as: $P(S \mid V) = \dfrac{P(S \cap V)}{P(V)}$

B. Interchanging labels and equivalent expressions

1. Since $P(S \cap V)$ is equivalent to $P(V \cap S)$, $P(S \cap V) = P(V) \times P(S \mid V)$
2. is equivalent to: $P(V \cap S) = P(S) \times P(V \mid S)$

C. And since, $P(S \mid V) = \dfrac{P(S \cap V)}{P(V)}$

1. substitution of $P(S) \times P(V \mid S)$ for $P(S \cap V)$
2. gives the expression,

$$P(S \mid V) = \frac{P(S) \times P(V \mid S)}{P(V)}$$

3. Which is **Bayes' Theorem**

VII. **BAYES' THEOREM**

A. General form of the habitat model: $P(S \mid E) = \dfrac{P(S)\,P(E \mid S)}{P(E)}$

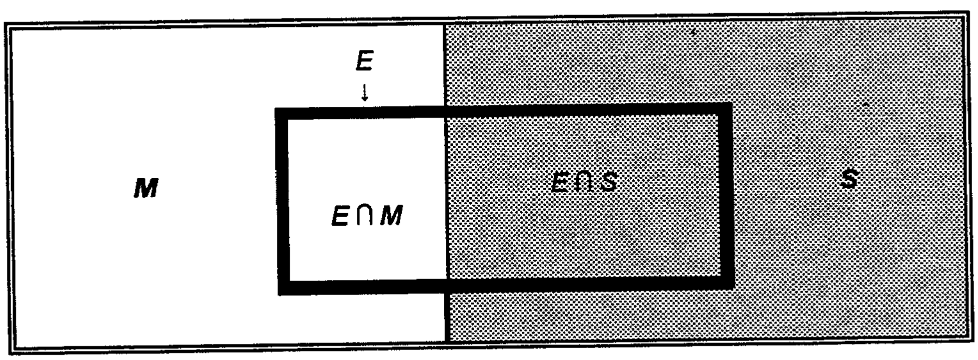
1. Where $P(S \mid E)$ is the **posterior probability** of **suitable habitat** (S) *given* **environmental conditions** (E)
2. $P(S)$ is the **prior probability** of **suitable habitat** (S)
3. $P(E \mid S)$ is the **conditional probability** of **environmental condition** (E) *given* the presence of **suitable habitat** (S)
4. $P(E)$ is the **total probability** that **environmental conditions** (E) are present

VIII. **Law of Total Probability**

A. The **total probability** of E draws on concepts of **union** and **intersecting** events

$$P(E) = P(S)\,P(E \mid S) + P(M)\,P(E \mid M)$$

B. Venn Diagram illustrating this relationship



C. So that **Bayes' Theorem** can be written as

$$P(S \mid E) = \frac{P(S)\,P(E \mid S)}{P(S)\,P(E \mid S) + P(M)\,P(E \mid M)}$$

IX. Bayesian Probability Models for Wildlife Habitat

A. Classify the habitat into one of two or more mutually exclusive categories:
SUITABLE HABITAT vs MARGINAL HABITAT

B. Estimate the **prior probability** that a parcel falls into each category:
   1. $P(S)$ = **prior probability** that the parcel is **suitable habitat**
   2. $P(M)$ = **prior probability** that the parcel is **marginal habitat**

C. Identify significant habitat attributes (E) that determine habitat as **suitable** or **marginal**.

D. For each attribute (E) and in the absence of local, site-specific animal-use frequencies of habitat attributes, develop estimates of the **conditional probability** that, given a particular habitat category (**suitable** or **marginal**), the attribute (E) is present from published literature.
   1. $P(E \mid S)$ = **conditional probability** of the attribute (E) being present given that the area is **suitable habitat**
   2. $P(E \mid M)$ = **conditional probability** of the attribute (E) being present given that the area is **marginal habitat**

E. Calculate the probabilities that the observed **set** of habitat attributes would be present, given that the parcel is within a specific habitat category (**suitable** or **marginal** habitat). These **posterior probabilities** are derived from the **prior probabilities** the **set** of **conditional probabilities**.
   1. $P(S \mid E)$ = the **posterior probability** of **suitable** habitat given the **set** of habitat attributes (E), and
   2. $P(M \mid E)$ = the **posterior probability** of **marginal** habitat given the **set** of habitat attributes (E)

G. Use **Bayes' Theorem** to calculate **posterior probabilities**:

1. $$P(S \mid E) = \frac{P(S) \times P(E \mid S)}{P(S) \times P(E \mid S) + P(M) \times P(E \mid M)}$$

2. $$P(M \mid E) = \frac{P(M) \times P(E \mid M)}{P(S) \times P(E \mid S) + P(M) \times P(E \mid M)}$$