Probability and Inference

Shiwali Mohan

Computer Science and Engineering
University of Michigan

May 25, 2012

Outline

1. Probability review
2. Probability and Inference from Full Joint Distribution
3. Bayes Net
4. Bayes Net

Terminology

- Prior probability

Terminology

- Prior probability
  - belief in propositions in the absence of any other information
  - marginal, unconditional

Terminology

- Prior probability
- belief in propositions in the absence of any other information
- marginal, unconditional
  - P(double): ?

Terminology

- Prior probability
- belief in propositions in the absence of any other information
- marginal, unconditional
- P(double): ?
- Posterior Probability

Terminology

- Prior probability
- belief in propositions in the absence of any other information
- marginal, unconditional
- P(double): ?
- Posterior Probability
- belief in propositions after evidence
- conditional
- Posterior probability belief in propositions after evidence P(double/5):? P(11/6):?

Terminology

- Prior probability
- belief in propositions in the absence of any other information
- marginal, unconditional
- P(double): ?

- Posterior Probability
- belief in propositions after evidence
- conditional
- P(double/5):?
- Posterior probability belief in propositions after evidence P(double/5):? P(11/6):?

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?
- hairy palms (palms), howl (howls), werewolf (werewolf)

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?
- hairy palms (palms), howl (howls), werewolf (werewolf)
- What is the bound on probability that a randomly selected person from the population howls at the moon?

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?
- hairy palms (palms), howl (howls), werewolf (werewolf)
- What is the bound on probability that a randomly selected person from the population howls at the moon?
- $0.05 <= P(\text{howls}) <= 0.25$

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?
- hairy palms (palms), howl (howls), werewolf (werewolf)
- What is the bound on probability that a randomly selected person from the population howls at the moon?
- 0.05 <= P(howls) <= 0.25
- If one in hundred people is a werewolf, what is the P(howls)?

Calculating Probabilities

Dr. Lycan Thropes, an expert in werewolves, asserts that a person is a werewolf iff they have hairy palms and howl at the moon. Further, he states that 25% of the population either has hairy palms or howls or both, and that 20% of the population has hairy palms.

- Random variables?
- hairy palms (palms), howl (howls), werewolf (werewolf)
- What is the bound on probability that a randomly selected person from the population howls at the moon?
- 0.05 $<=$ P(howls) $<=$ 0.25
- If one in hundred people is a werewolf, what is the P(howls)?
- 0.06

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

- What is the marginal probability of werewolf?

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

- What is the marginal probability of werewolf?
- 0.045

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

- What is the marginal probability of werewolf?
- 0.045
- If an individual has hairy palms and howls, what is the probability that the individual is a werewolf?

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

- What is the marginal probability of werewolf?
- 0.045
- If an individual has hairy palms and howls, what is the probability that the individual is a werewolf?

Inference on Full Joint

Noted werewolf specialist Prof. Loupgarou disagrees with Thropes, and provides a joint probability table for combinations of werewolves, hairy palms, and howling.

- What is a full joint distribution?
- How many entries exist in the full joint distribution?
- Given the following full joint distribution. Sum to 1?

hairy palms	howls	werewolf	Probability
false	false	false	0.7
false	false	true	0.005
false	true	false	0.05
false	true	true	0.01
True	false	false	0.2
True	false	true	0.01
True	true	false	0.005
True	true	true	0.02

- What is the marginal probability of werewolf?
- 0.045
- If an individual has hairy palms and howls, what is the probability that the individual is a werewolf?

Inference

Events A, B, C, D, E occur with the following probabilities. Assume conditional independence amongst variables.

```
P(A) = 0.3
P(B) = 0.6
P(C|A) = 0.8
P(C|-A) = 0.4
P(D|A,B) = 0.7
P(D|A,-B) = 0.8
P(D|-A,B) = 0.1
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|C) = 0.2
```

For given probabilities
P(A) = 0.3 P(B) = 0.6 P(C|A) = 0.8 P(C|-A) = 0.4 P(D|A,B) = 0.7 P(D|A,-B) = 0.8 P(D|-A,B) = 0.1 P(D|-A,B) = 0.2 P(E|C) = 0.7 P(E|C) = 0.2

- Construct a Bayes Net

- P(D)?

  P(D,A,B) + P(D,A,-B) + P(D,-A,B) + P(D,-A,-B) = P(D|A,B) P(A,B) + P(D|A,-B) P(A,-B) + P(D|-A,B) P(-A,B) + P(D|-A,-B) P(-A,-B) = (since A and B are independent absolutely) P(D|A,B) P(A) P(B) + P(D|A,-B) P(A) P(-B) + P(D|-A,B) P(-A) P(B) + P(D|-A,B) P(-A) P(-B) = 0.7*0.3*0.6 + 0.8*0.3*0.4 + 0.1*0.7*0.6 + 0.2*0.7*0.4 = 0.32

- P(A|C)?

  P(A|C) = P(C|A)P(A) / P(C). Now P(C) = P(C,A) + P(C,-A) = P(C|A)P(A) + P(C|-A)P(-A) = 0.8*0.3+ 0.4*0.7 = 0.52 So P(C|A)P(A) / P(C) = 0.8*0.3/0.52= 0.46.

- P(C|-A,E)?

  P(C|-A,E) = P(E|C,-A) * P(C|-A) / P(E|-A) = P(E|C) * P(C|-A) / P(E|-A). Now P(E|-A) = P(E,C|-A) + P(E,-C|-A) = P(E|C,-A) P(C|-A) +

P(E|-C,-A) P(-C|-A) = (since E is independent of A given C) P(E|C) * P(C|-A) + P(E|-C) * P(-C|-A). So we have P(C|-A, E) = P(E|C) * P(C|-A) / (P(E|C) * P(C|-A) + P(E|-C) * P(-C|-A)) = 0.7*0.4 / (0.7 * 0.4 + 0.2 * 0.6) = 0.7

```
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|-C) = 0.2
```

Inference

For given probabilities

```
P(A) = 0.3
P(B) = 0.6
P(C|A) = 0.8
P(C|-A) = 0.4
P(D|A,B) = 0.7
P(D|A,-B) = 0.8
P(D|-A,B) = 0.1
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|-C) = 0.2
```

- P(A|C)?

Inference

For given probabilities

```
P(A) = 0.3
P(B) = 0.6
P(C|A) = 0.8
P(C|-A) = 0.4
P(D|A,B) = 0.7
P(D|A,-B) = 0.8
P(D|-A,B) = 0.1
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|-C) = 0.2
```

- $P(A|C)$?

- ```
  P(A|C) = P(C|A)P(A) / P(C)
  P(C) = P(C,A) + P(C,-A)
       = P(C|A)P(A) + P(C|-A)P(-A)
       = 0.8*0.3+ 0.4*0.7 = 0.52
  P(C|A)P(A) / P(C) = 0.8*0.3/0.52= 0.46.
  ```
  ```

Inference

For given probabilities

```
P(A) = 0.3
P(B) = 0.6
P(C|A) = 0.8
P(C|-A) = 0.4
P(D|A,B) = 0.7
P(D|A,-B) = 0.8
P(D|-A,B) = 0.1
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|-C) = 0.2
```

- P(C|-A,E)?

Inference

For given probabilities

```
P(A) = 0.3
P(B) = 0.6
P(C|A) = 0.8
P(C|-A) = 0.4
P(D|A,B) = 0.7
P(D|A,-B) = 0.8
P(D|-A,B) = 0.1
P(D|-A,-B) = 0.2
P(E|C) = 0.7
P(E|-C) = 0.2
```

- P(C|-A,E)?

-
```
P(C|-A,E)
= P(E|C,-A) * P(C|-A) / P(E|-A)
= P(E|C) * P(C|-A) / P(E|-A).
P(E|-A) = P(E,C|-A) + P(E,-C|-A)
        = P(E|C,-A) P(C|-A) + P(E|-C,-A) P(-C|-A)
          (since E is independent of A given C)
         = P(E|C) * P(C|-A) + P(E|-C) * P(-C|-A)
P(C|-A, E)
= P(E|C) * P(C|-A) / (P(E|C) * P(C|-A) + P(E|-C) * P(-C|-A))
```