Lecture 13: Additional bits on Conditional Probability and Independence

13.1: Independence of multiple events (Ross, 3.4)

(i) Pairwise independence (reminder)

Recall E and F are independent if $P(E \cap F) \;=\; P(E) \times P(F)$.

Then $P(E \mid F) \;\equiv\; P(E \cap F)/P(F) \;=\; P(E)$ and $P(F \mid E) \;\equiv\; P(E \cap F)/P(E) \;=\; P(F)$.

Recall that if E and F are independent, so are E and F^c, E^c and F, and E^c and F^c.

(ii) Joint independence

E_1, E_2, ..., E_n are (jointly) independent if for every subset $E_{r_1}, E_{r_2}, ...$ with $r_1 < r_2 < .... \leq n$

$$P(E_{r_1} \cap E_{r_2} \cap .... \cap r_k) \;=\; P(E_{r_1}) \times P(E_{r_2}) \times .... \times P(E_{r_k}).$$

(iii) Pairwise independence without joint independence: example.

Two independent rolls of a fair die. D_1 is first throw gives odd number.

D_2 is second throw gives odd number. D_3 is sum of two throws is odd number.

$P(D_1) \;=\; P(D_2) \;=\; P(D_3) \;=\; 1/2.$ $\qquad P(D_1 \cap D_2) \;=\; P(D_1 \cap D_3) \;=\; P(D_2 \cap D_3) \;=\; 1/4.$

But $P(D_1 \cap D_2 \cap D_3) \;=\; 0$, not 1/8. These three events are pairwise independent but NOT jointly independent.

(iv) The three-way independence, without pairwise, is clearly also possible.

Let F, G, I be Swiss adults fluent in French, German and Italian.

Suppose $P(F) \;=\; P(G) \;=\; P(I) \;=\; 1/2$, and $P(F \cap G \cap I) \;=\; P(F) \times P(G) \times P(I) \;=\; 1/8$.

But this does not determine $P(F \cap G)$ etc.

For example, we could have as shown. Then

$P(F \cap G) \;=\; 3/8, \;\; P(F \cap I) \;=\; 1/8, P(G \cap I) = 1/8.$

	I			I^c		
	F	F^c		F	F^c	
G	1/8	0		1/4	1/8	1/2
G^c	0	3/8		1/8	0	1/2
	1/8	3/8		3/8	1/8	

13.2: Conditional dependence for random variables

(i) Conditioning a discrete random variable

Recall $P(X = x)$ is just an event, and $P(X \in B) = \sum_{x \in B} P(X = x)$.

So if X is Bin(10,0.5): $P(X \text{ even}) \;=\; 1/2$ and $P(X = x \mid X \text{ even}) \;=\; 2P(X = x)$ if x is even, 0 otherwise.

If X is Poisson, parameter 1: $P(X \geq 2) \;=\; 1 - e^{-1} - e^{-1}$, and $P(X = x \mid X \geq 2) \;=\; (e^{-1}/x!)/(1 - 2e^{-1})$.

(ii) Conditioning a continuous random variable

Recall $X \in B$ is an event and $P(X \in B) = \sum_{x \in B} P(X = x)$.

So $P(X \in C | X \in B) \;=\; P(X \in B \cap C)/P(X \in B) \;=\; \int_{B \cap C} f(x)dx / \int_B f(x)dx$

13.3 Examples of conditioning a continuous random variable

(i) Example of a Uniform random variable Suppose X has p.d.f. $f(x) = 1$, $0 \leq x \leq 1$.

$P(X > 0.6 \mid X \leq 0.8) \;=\; P(0.6 < X \leq 0.8)/P(X \leq 0.8) \;=\; 0.2/0.8 \;=\; 0.25.$

(ii) Example for an exponential random variable

Suppose X has p.d.f. $f(x) = 0.5e^{-0.5x}$ on $0 < x < \infty$: $\int f(x)dx \;=\; -e^{-0.5x}$.

So $P(X \leq 6 \mid X > 2) = P(2 < X \leq 6)/P(X > 2) \;=\; (e^{-1} - e^{-3})/e^{-1} \;=\; (1 - e^{-2}) \;\approx\; 6/7.$

(iii) The **forgetting property** of the exponential.

Suppose X has p.d.f. $f(x) = \lambda \exp(-\lambda x)$, $\;0 < x < \infty$. Consider

$P(X > a + b \mid X > a) = P(X > a + b)/P(X > a) = \exp(-\lambda(a + b))/\exp(-\lambda a) = \exp(-\lambda b) = P(X > b).$

Lecture 14: Examples of independence and conditioning

14.1 Independent inheritance of characteristics: no genetic linkage

To find the flower-color of offspring pea-plants, Mendel had to grow them. Things he could observe in the seeds were much easier to study.

Unfortunately, he found that the DNA determining these characters is inherited *independently* of the DNA determining flower-color. One such character is seed-coat color: green (G) or yellow (Y). Plants that are GG or GY have green seeds, ones that are YY have yellow seeds.

(i) Mendel crossed two red-flowered pea-plants whose parents were a red-flowered plant and a white-flowered plant. What is the proportion of red-flowered offspring? and of white?

(ii) Mendel crossed two green-seed pea-plants whose parents were a green-seed plant and a yellow-seed plant. What is the proportion of green-seed offspring? and of yellow?

(iii) Mendel crossed two red-flowered, green-seed, pea plants, whose parents were red-flowered/green-seed and white-flowered/yellow-seed. In the offspring, what is the proportion of red-green, red-yellow, white-green, and white-yellow?

(iv) Mendel crossed two red-flowered, green-seed, pea plants, whose parents were red-flowered/yellow-seed and white-flowered/green-seed. In the offspring, what is the proportion of red-green, red-yellow, white-green, and white-yellow?

14.2 Approximating discrete random variables

(i) Back to the bacteria (or cells) of lecture 11. Let T be time (continuous) or number of generations (continuous) back to to common ancestor.

Model 1: At each generation back, two bacteria share a parent with probability $p = 0.02$, independently at each generation.

Model 2: The time back to a common ancestor is exponential with rate parameter 0.02.

(a) Under model 1, what is $P(T > 50)$?

(b) Under model 2, what is $P(T > 50)$?

(ii) Repeat (i), for $p = 0.001$ and $P(T > 1000)$ for model 1 and model 2.

14.3 Conditioning continuous random variables

(i) Using model 2 of 14.2 (i), find $P(T > 60 \mid T > 10)$.

What about $P(T > 550 \mid T > 500)$?

(ii) Using the unrealistic model that T has a p.d.f. uniform on (0,100), find $P(T \leq 25 \mid T \leq 50)$.

Compare this with $P(T \leq 25)$.

(iii) Using model 2 of 14.2 (i), find $P(T \leq 25 \mid T \leq 50)$.

Compare this with $P(T \leq 25)$.

Lecture 15: Mean and variance of random variables: Ross 4.3-4.5, 5.2

15.1: Expected value of a random variable: the mean

(i) Discrete case (Ross, 4.3)

If X is discrete with p.m.f. $P(X = x) = p(x) > 0$ for $x \in \mathcal{X}$, the *expected value* of X denoted $\mathrm{E}(X)$ is $\mathrm{E}(X) = \sum_{x \in \mathcal{X}} x\, p(x)$, provided this sum exists and is finite.

(ii) Continuous case (Ross 5.2)

If X is continuous with p.d.f. $f(x)$, $f(x) \geq 0$ for $-\infty < x < \infty$. the *expected value* of X denoted $\mathrm{E}(X)$ is $\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx$, provided this integral exists and is finite. (Note $f(x)\, dx \approx P(x < X \leq x + dx)$.)

15.2: Expected value of a function of a random variable

(i) Discrete case (Ross 4.4)

Now if X takes values x_i, the values taken by $Y = g(X)$ are $g(x_i)$, but several x_i may have the same $g(x_i)$.

$$P(Y = y) \;=\; \sum_{i:g(x_i)=y} p(x_i) \quad \text{so}$$

$$\mathrm{E}(g(X)) \;=\; \sum_y y\; P(Y=y) \;=\; \sum_y \left(y \sum_{i:g(x_i)=y} p(x_i) \right) \;=\; \sum_y \sum_{i:g(x_i)=y} g(x_i) p(x_i) \;=\; \sum_i g(x_i) p(x_i).$$

(ii) A simple property (Ross, Corollary 4.1)

$$\mathrm{E}(aX+b) \;=\; \sum_x (ax+b)p(x) \;=\; a \sum_x x\, p(x) \;+\; b \sum_x p(x) \;=\; a\, \mathrm{E}(X) \;+\; b$$

The same result holds for continuous random variables.

(iii) A neat result for non-negative random variables (Ross 5.2, Lemma 2.1)

$$\mathrm{E}(X) \;=\; \int_{y=0}^{\infty} y\; f(y)\; dy \;=\; \int_{y=0}^{\infty} \left(\int_{x=0}^{y} dx \right) f(y)\; dy \;=\; \int_{x=0}^{\infty} \left(\int_{y=x}^{\infty} f(y)\; dy \right) dx \;=\; \int_{x=0}^{\infty} P(X > x) dx$$

(iv) Continuous case (Ross 5.2: proposition 2.1).

For a continuous random variable $\mathrm{E}(g(x) \;=\; \int_x g(x)\; f(x)\; dx$.

We can prove this for a non-negative $g(X)$, using (iii): see Ross 5.2 for details.

15.3: The variance of a random variable (Ross 4.5, 5.2)

(i) Definition: If $E(X) = \mu$, $\mathrm{var}(X) \;=\; \mathrm{E}(X-\mu)^2$

Since $(x-\mu)^2 \geq 0$ for every x, the definition shows $\mathrm{var}(X) \;\geq\; 0$.

(ii) Property 1: using 15.2 (ii) we have

$$\mathrm{var}(X) \;=\; \mathrm{E}(X-\mu)^2 \;=\; \mathrm{E}(X^2 - 2\mu X \;+\; \mu^2) \;=\; \mathrm{E}(X^2) - 2\mu \mathrm{E}(X) \;+\; \mu^2 \;=\; \mathrm{E}(X^2) - 2\mu \times \mu \;+\; \mu^2$$
$$\;=\; \mathrm{E}(X^2) - \mu^2 \;=\; \mathrm{E}(X^2) - (\mathrm{E}(X))^2.$$ (This is usually the easiest way to compute $\mathrm{var}(X)$.)

(iii) Property 2: using 15.2 (ii) we have

$$\mathrm{var}(aX+b) \;=\; \mathrm{E}((aX+b-a\mu-b)^2) \;=\; \mathrm{E}(a^2(X-\mu)^2) \;=\; a^2\mathrm{E}((X-\mu)^2) \;=\; a^2\mathrm{var}(X).$$

15.4 One example: the mean and variance of a Poisson random variable $P(X = x) = e^{-\lambda}\lambda^x/x!$.

The mean: $\mathrm{E}(X) \;=\; \sum_{x=0}^{\infty} x e^{-\lambda}\lambda^x/x! \;=\; \sum_{x=1}^{\infty} x e^{-\lambda}\lambda^x/x! \;=\; e^{-\lambda}\lambda \sum_{x=1}^{\infty} \lambda^{x-1}/(x-1)! \;=\; e^{-\lambda}\lambda e^{\lambda} \;=\; \lambda$

Then: $\mathrm{E}(X(X-1)) \;=\; \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\lambda^x/x! \;=\; e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \lambda^{x-2}/(x-2)! \;=\; e^{-\lambda}\lambda^2 e^{\lambda} \;=\; \lambda^2$

So then $\mathrm{var}(X) \;=\; \mathrm{E}(X^2) - (\mathrm{E}(X))^2 \;=\; \mathrm{E}(X(X-1)) \;+\; \mathrm{E}(X) - (\mathrm{E}(X))^2 \;=\; \lambda^2 + \lambda - \lambda^2 \;=\; \lambda$