

## The quantitative rules

### Proof of 2.19

Credit to atwwgb

Continuing to get 2.19 from 2.18:

We have then  $G(x, y)G(y, z) = P(x, z)$ . Pick any fixed  $z$ . Denote  $P(x, z) = A(x)$  and  $G(y, z) = B(y)$ . Then  $G(x, y) = \frac{A(x)}{B(y)}$  [and  $G(y, z) = \frac{A(y)}{B(z)}$ ].

Plug this in to  $G(x, y)G(y, z) = P(x, z)$  to get  $\frac{A(x)A(y)}{B(y)B(z)} = P(x, z)$ . So  $A(y)/B(y)$  is independent of  $y$ , so is constant equal to  $r$ . This means

$$G(x, y) = \frac{A(x)}{B(y)} = \frac{A(x)A(y)}{A(y)B(y)} = r \frac{A(x)}{A(y)}$$

### Proof of Equation 2.50

Source: stackexchange, I've reworded it and added detail to (hopefully) make it clearer.

We will use the Taylor series approximation, which is an approximation of  $f(t)$  around the point  $a$ :

$$f(t) = f(a) + f'(a)(t - a) + O((t - a)^2)$$

Big O notation is described on Wikipedia.

The proof:

Letting  $\delta = e^{-q}$ , we have from (2.48):

$$S(y) = S\left[\frac{S(x)}{1 - \delta}\right]$$

We then use a Taylor series approximation of the function  $f(\delta) = \frac{1}{1 - \delta}$  around with  $a = 0$ .

$$S(y) = S[S(x)(1 + \delta + O(\delta^2))]$$

$$S(y) = S[S(x) + S(x)\delta + S(x)O(\delta^2)]$$

Now we want to get rid of the  $S[]$  surrounding the equation, so we will use another Taylor approximation of the function  $S(t)$ . We approximate around the point  $a = S(x)$ .

This gives us the approximation of  $S(t)$  as:

$$S(t) = S[S(x)] + S'[S(x)](t - S(x)) + O((t - S(x))^2)$$

Letting  $t = S(x) + S(x)\delta + S(x)O(\delta^2)$

$$S[S(x) + S(x)\delta + S(x)O(\delta^2)] = S[S(x)] + S'[S(x)](S(x)\delta + S(x)O(\delta^2)) + O((S(x)\delta + S(x)O(\delta^2))^2)$$

$$S[S(x) + S(x)\delta + S(x)O(\delta^2)] = S[S(x)] + S'[S(x)]S(x)\delta + S'[S(x)]S(x)O(\delta^2) + O((S(x)\delta + S(x)O(\delta^2))^2)$$

With big O notation we can get rid of constant factors:

$$S[S(x) + S(x)\delta + S(x)O(\delta^2)] = S[S(x)] + S'[S(x)]S(x)\delta + O(\delta^2) + O((\delta + O(\delta^2))^2)$$

With big O notation we can also get rid terms that drop asymptotically faster than the largest term.

$$S[S(x) + S(x)\delta + S(x)O(\delta^2)] = S[S(x)] + S'[S(x)]S(x)\delta + O(\delta^2)$$

□

## Explanation of 2.52

TODO

## Exercise 2.1

I think this problem is ambiguous and can be interpreted in multiple ways, see here for a different interpretation. But I think the following interpretation makes more sense.

With  $X$  representing any background information:

$$\begin{aligned} p(C|(A+B)X) &= \frac{p(A+B|CX)p(C|X)}{p(A+B|X)} \\ &= \frac{(p(A|CX) + p(B|CX) - p(AB|CX))p(C|X)}{p(A|X) + p(B|X) - p(AB|X)} \\ &= \frac{p(AC|X) + p(BC|X) - p(ABC|X)}{p(A|X) + p(B|X) - p(AB|X)} \end{aligned}$$

### **Exercise 2.2**

TODO

### **Exercise 2.3**

TODO