LAW OF TOTAL EXPECTATION

$$E(\underbrace{E(Y \mid X)}) = E(Y)$$

outside: expected value of a function of X

inside: conditional expectation with respect to the distribution of Y (X "fixed")

example: X and Y as on Friday: $f(x,y) = \frac{1}{2}$ for $0 \le x \le y \le 2$.

First: $E(Y \mid X = x_0) = \frac{2+x_0}{2}$, equivalently, $E(Y \mid X) = \frac{2+X}{2}$

a function of the rv X

Then: $E(E(Y \mid X)) = E\left(\frac{2+X}{2}\right) = \int_0^2 \frac{2+x}{2} f_X(x)\,dx = \int_0^2 \frac{2+x}{2} \cdot \frac{2-x}{2}\,dx = \boxed{\frac{4}{3}}$

Find $E(Y)$ from $f_Y(y) = \int_0^y \frac{1}{2}\,dx = \frac{y}{2}$ for $0 \le y \le 2$

$E(Y) = \int_0^2 y \frac{y}{2}\,dy = \int_0^2 \frac{y^2}{2}\,dy = \boxed{\frac{4}{3}}$

Same!

LAW OF TOTAL VARIANCE

$$Var(Y) = Var(E(Y \mid X)) + E(Var(Y \mid X))$$

0. The number of eggs N found in a nest of a certain species of turtle has a Poisson distribution with mean λ. Each egg has a probability p of being viable, and this event is independent from egg to egg. Find the mean and variance of the number of viable eggs per nest.

rvs: $N \sim \text{Poisson}(\lambda)$

Y = number of viable eggs $\sim Bin(N, p)$

mean of Y: $E(Y) = E(E(Y \mid N)) = E(Np) = p\,E(N) = p\lambda$

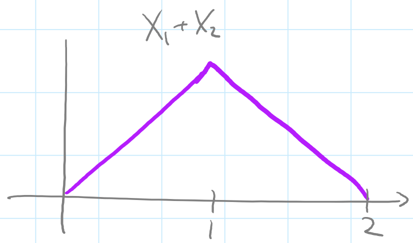
Law of total expectation

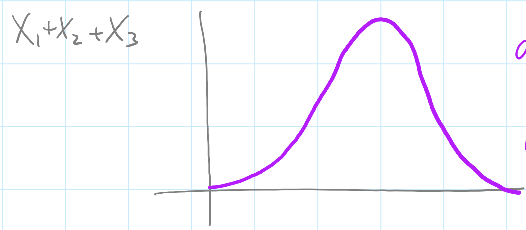
Variance:
$$\begin{aligned} Var(Y) &= Var(E(Y\mid N)) + E(Var(Y\mid N)) \\ &= Var(Np) + E(Np(1-p)) \\ &= p^2\, Var(N) + p(1-p)\,E(N) \\ &= p^2\lambda + p(1-p)\lambda = \cancel{p^2\lambda} + p\lambda - \cancel{p^2\lambda} = p\lambda \end{aligned}$$

---

1. Sketch the density of the sum of two iid Unif[0,1] random variables. How about three? Four? More?

independent identically distributed





density of the sum $X_1 + \cdots + X_n$ looks more and more normal as $n \to \infty$.

2. Use **R** to simulate 10,000 samples from the distribution Unif[0,1]:

```
hist(runif(10000))
```

What is the shape of the histogram?

approximately uniform

3. Now simulate 10,000 averages, each of 30 samples from Unif[0,1]:

```
means <- replicate( 10000, mean(runif(30)) )
hist(means, breaks=30)
```

What is the shape of the histogram?

approximately normal!

That is: if Y is the sum of 30 iid Unif[0,1] random variables, then the distribution of Y is approximately normal.

4. Repeat #2 and #3 above with a different distribution. That is, replace `runif` with a distribution of your choice. What do you observe?

If Y is the sum of 30 iid random variables from any of the distributions we have studied, then the distribution of Y is approximately normal.