$$E[g(x)] = \int g(x) f_X(x)\, dx. \qquad \text{LOTUS.}$$

$X \sim f_X(x).$

$Y = g(x).$

$$Y \sim \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$$

<u>Proof.</u>

find CDF of Y, then take derivatives

$$P(Y \leq y) = P(g(X) \leq y)$$
$$= P(X \leq g^{-1}(y))$$
$$= F_X(g^{-1}(y)) = F_X(x).$$

take derivatives wrt y.

$$f_Y(y) = f_X(x) \frac{dx}{dy}.$$

$$f_X(x)\, dx = f_Y(y)\, dy.$$

$T = X + Y$ $\qquad$ $E[T] = E[X] + E[Y]$

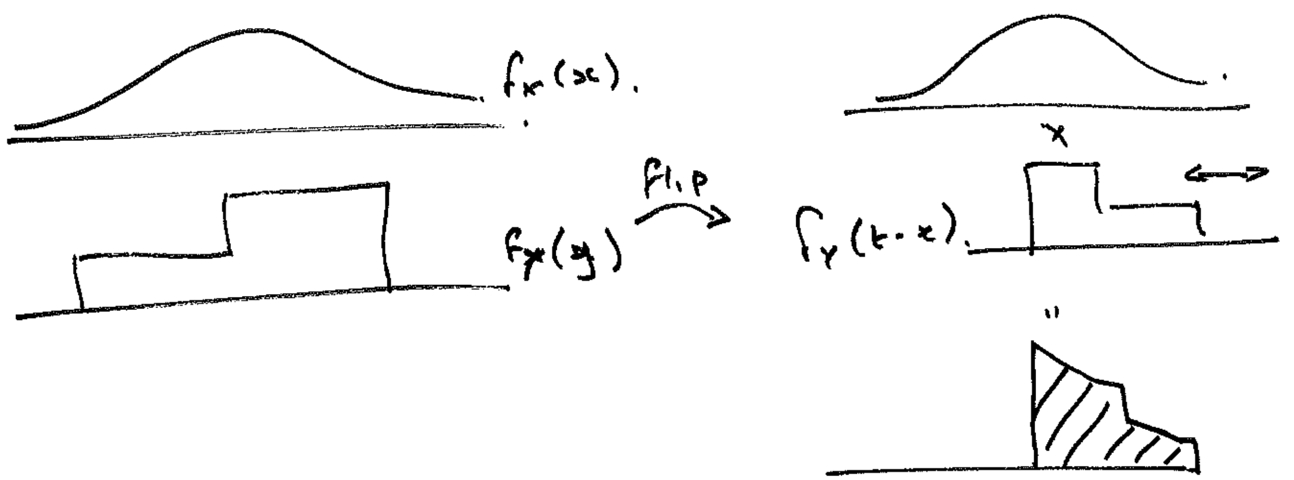
$f_T(t) = ?$ $\qquad$ Assume X, Y independent.

discrete case

$$P(T = t) = \sum_{x=0}^{t} P(X = x)\, P(Y = t - x)$$

continuous

$$f_t(t) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(t - x)\, dx \qquad \text{convolution.}$$



CDF $\quad F_T(t) = P(T \leq t)$

$$= \int_{-\infty}^{\infty} P(X + Y \leq t \mid X = x) f_X(x)\, dx$$

LOTP.

$$= \int_{-\infty}^{\infty} P(Y \leq t - x) f_X(x)\, dx$$

substitute $X = x$ and drop the conditioning as X, Y are independ

$$F_T(t) = \int_{-\infty}^{\infty} F_Y(t-x) f_X(x)\, dx.$$

take derivatives w.r.t t

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)\, dx.$$

Example $\quad X \sim N(0,1)$
$\quad Y \sim N(0,1)$ $\quad$ independent.

$$T = X + Y$$

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t-x)^2}{2}\right)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[x^2 + t^2 - 2tx + x^2\right]\right)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[2x^2 - 2tx + t^2\right]\right) dx$$

$$2x^2 - 2tx + t^2$$

$$2\left(x^2 - tx + \frac{t^2}{2}\right).$$

$$2\left(\left(x - \frac{t}{2}\right)^2 - \frac{t^2}{4} + \frac{t^2}{2}\right)$$

$$2\left[\left(x - \frac{t}{2}\right)^2 + \frac{t^2}{4}\right]$$

$$f_F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \times 2\left[\left(x - \frac{t}{2}\right)^2 + \frac{t^2}{4}\right]\right) \text{c}$$

$$= \frac{1}{2\pi} \exp\left(-\frac{t^2}{4}\right) \underbrace{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \times 2\left(x - \frac{t}{2}\right)^2\right) \text{c}}_{\text{this looks like the}}$$

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\,\frac{(x-\mu)^2}{\sigma^2}\right) dx = \sigma\sqrt{2\pi}$$

$$\mu = \frac{t}{2}$$

$$\sigma^2 = \frac{1}{2}$$

So $\displaystyle\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \times 2 \left(x - \frac{t}{2}\right)^2\right) dx$

$$= \frac{1}{\sqrt{2}}\sqrt{2\pi} = \sqrt{\pi}$$

$$f_T(t) = \frac{1}{2\pi} \exp\left(-\frac{t^2}{4}\right) \times \frac{\sqrt{2\pi}}{\sqrt{2}}$$

$$= \frac{1}{\sqrt{2}\,\sqrt{2\pi}} \exp\left(-\frac{1}{2}\,\frac{t^2}{2}\right) = N(0,2)$$

<u>Hence</u> the sum of two independent $N(0,1)$ RVs is an $N(0,2)$ RV.

In general, any linear combination of Normal RVs will be Normal.

$$\left\{\begin{array}{l} X \sim N(\mu_1, \sigma_1^2) \\ Y \sim N(\mu_2, \sigma_2^2) \end{array}\right. \Rightarrow X + Y \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

ariances.

oisson.

$X \sim \text{Poiss}(\lambda)$

Mean $= \lambda$

variance $= \lambda$

$$E[X^2] = \sum_{k=0}^{\infty} k^2 \, P(X=k)$$

$$= \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!}$$

$$\vdots$$

left as exercise the reader

$$= \lambda^2 + \lambda$$

Binomial.

$X \sim \text{Bin}(n, p)$

$$E[X^2] = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}.$$

X is sum of n iid Bernoulli (p) RVs

Var of a sum of independent RVs = sum of variances of the RVs.

(Proof later

$$X = X_1 + X_2 + X_3 + \dots X_n$$

$$= I_1 + I_2 + I_3 + \dots I_n \qquad I_i \sim \text{iid Bernoulli}$$

$$X^2 = (I_1 + I_2 + \dots I_n)(I_1 + I_2 + \dots I_n)$$

$$= I_1^2 + I_2^2 + \dots I_n^2$$

$$+ 2 I_1 I_2 + 2 I_1 I_3 + \dots + 2 I_{n-1} I_n$$

$$E[X^2] = n\, E[I_1^2] + 2\binom{n}{2} E[I_1 I_2]$$

each I_j^2 has the same distribution

$\binom{n}{2}$ cross terms.

I_1 - indicator variable of success on 1st trial.

I_1^2 - takes the values 0, 1 with the same probabilities that I_1 does

Expected value of an indicator RV

- prob. of the event that it is an indicator for.

$$E[I] = 0 \times P(I=0) + 1\, P(I=1)$$

$$\Rightarrow E[I_1^2] = p$$

$E[I_1 I_2]$

this is an indicator for <u>both</u> trials being successes

- which happens with prob. p^2

$$E[I_1 I_2] = p^2$$

$$E[X^2] = np + n(n-1)p^2$$

$$= np + n^2p^2 - np^2$$

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2$$

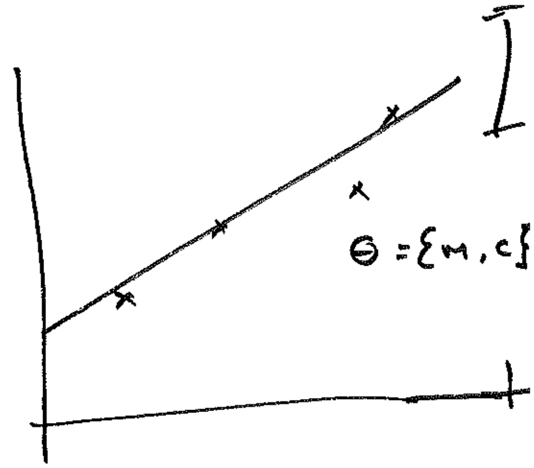
$$= np + \cancel{n^2p^2} - np^2 - \cancel{n^2p^2}$$

$$= np - np^2$$

$$= np(1-p)$$

$$= \underline{npq}.$$

hat can we predict about the
uture based on what we have
oserved so far?



———.

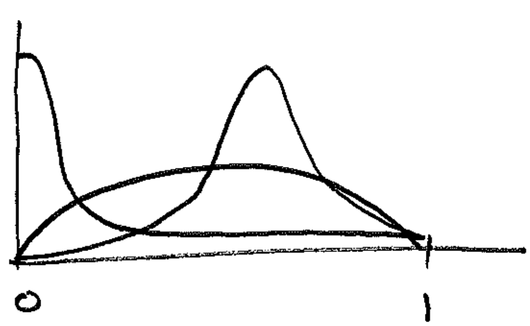
hat's the probability that
the sun will rise tomorrow?

un has risen for the last n days.

<u>lodel:</u> $X_1, X_2, \dots X_n \sim$ iid Bernoulli($\cancel{p}$ θ).

what i'm interested in is the
<u>unknown</u> success probability $\cancel{p}$ θ

└ θ is a Random Variable
it has a distribution $f_\Theta(\theta)$



$f_\Theta(\theta)$ encodes ~~one~~ our
knowledge of the success
prob. of the Bernoulli
distribution

$f_{\Theta|D}(\Theta = \theta \mid D = d)$

parameters of the model

data that we have observed.

$$f_{\Theta|D}(\theta \mid d) = \frac{f_{D|\Theta}(d \mid \theta)\, f_\Theta(\theta)}{f_D(d)}$$

Continuous fo[rm] of Bayes theorem

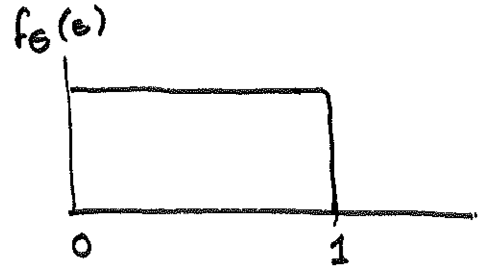
$$f_D(d) = \int f_{D|\Theta}(d \mid \theta) f_\Theta(\theta)\, d\theta \quad \text{LOTP.}$$

this is not a function of θ.

- Sometimes we ignore it, if we're interested in the <u>shape</u> of $f_{\Theta|d}($

- sometimes we can recognize the numerator as a distribution we know about, and so we know directly how it is normalized.

Before we collect data.



$U(0,1)$ – we have no idea what values of Θ (success probability) are likely before collecting data.

$f_{D|\Theta}(d \mid \theta)$

Assume a value for the success prob. Θ. What's the probability of observing the data?

n successes in n trials

$$p(\text{data} \mid \theta) = \theta^n$$

$$p(\theta \mid \text{data}) \propto \theta^n = c\,\theta^n$$

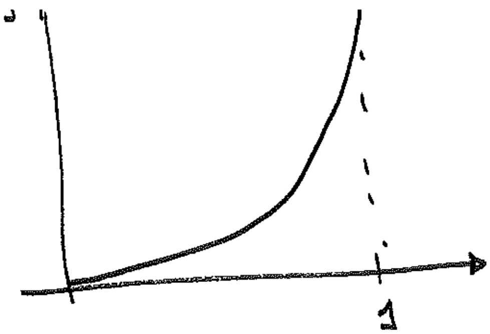
Hence $$\int_0^1 c\,\theta^n \, d\theta = 1.$$

$$c \left. \frac{\theta^{n+1}}{n+1} \right|_0^1 = 1$$

$$c = n+1$$

$$p(\theta \mid \text{data}) = (n+1)\,\theta^n$$

$$f_{\Theta|D}(\theta \mid d) = (n+1)\,\theta^n$$



This summarizes our knowledge of the success prob. after having observed the data.

What's the prob. that the sun will rise tomorrow?

- prob of success on $(n+1)^{th}$ trial, given data for 1^{st} n trials.

$$P(X_{n+1} = 1 \mid X_1, X_2 \ldots X_n)$$

$$= \int_0^1 P(X_{n+1} = 1 \mid \theta)\, f_{\Theta|D}(\theta \mid d)\, d\theta \qquad \underline{\text{LOTP.}}$$

$$= \int_0^1 \theta\,(n+1)\,\theta^n \, d\theta.$$

$$= \frac{n+1}{n+2}.$$

← <u>Predictive</u> probability.
(prob. of some unobserved events based on the event we have observed)