Probability

Given PDF $p(x)$, and random variable $F(x)$,

Expectation of $F(x)$: $\langle F(x)\rangle = \int_{-\infty}^{\infty} dx\, p(x) F(x)$.

By construction, PDF of $F(x)$ of some f s.t $f = F(x_i)$ for all possible x_i

then $p_F(f)\,df = \sum_{x_i} p(x_i)\,dx_i$

$p_F(f) = \sum_i p(x_i) \left|\frac{dx}{dF}\right|_{x=x_i}$ ie $F(x) = x^2$ $p(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ where $\left.\frac{dx}{dF}\right|_{x_i} = \left|\frac{1}{2\sqrt{F(x)}}\right|_{x_i}$; $x_i = \pm\sqrt{f}$

Moments of PDF $m_n = \langle x^n\rangle = \int dx\, p(x)\, x^n$

characteristic func. (generator of moments of distribution)

$$\boxed{\tilde{p}(k) = \langle e^{-ikx}\rangle} = \int dx\, p(x)\, e^{-ikx} = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!}\langle x^n\rangle$$

introduce cumulant generating func $\langle x^n\rangle_c$, $\boxed{\ln\tilde{p}(k) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!}\langle x^n\rangle_c}$ w/ $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} \dots$

coefficient of $\frac{(-ik)^n}{n!}$

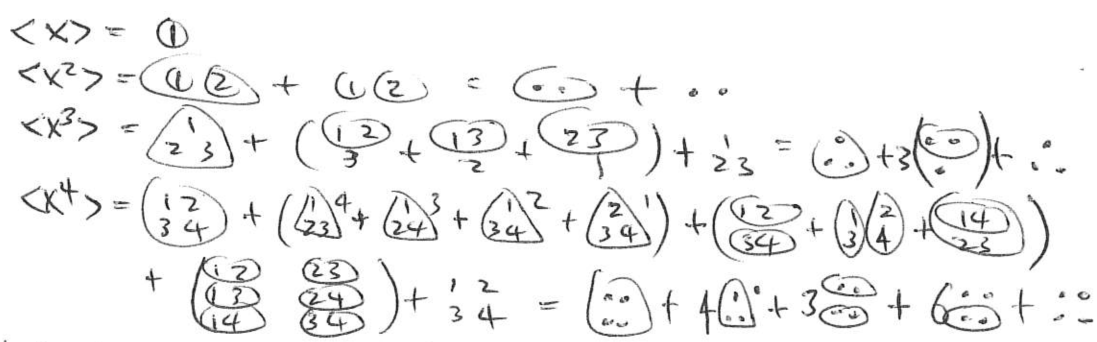
we have $\langle x\rangle_c = \langle x\rangle$

$\langle x^2\rangle_c = \langle x^2\rangle - \langle x\rangle^2$

$\langle x^3\rangle_c = \langle x^3\rangle - 3\langle x^2\rangle\langle x\rangle + 2\langle x\rangle^3$

$\vdots$

Graphical techniques



thus the corresponding algebraic expression.

$\langle x\rangle = \langle x\rangle_c$ $\langle x^3\rangle = \langle x^3\rangle_c + 3\langle x^2\rangle_c\langle x\rangle_c + \langle x\rangle_c^3$

$\langle x^2\rangle = \langle x^2\rangle_c + \langle x\rangle_c^2$ $\langle x^4\rangle = \langle x^4\rangle_c + 4\langle x^3\rangle_c\langle x\rangle_c + 3\langle x^2\rangle_c^2 + 6\langle x^2\rangle_c\langle x\rangle_c^2 + \langle x\rangle_c^4$

Algebraically. $\sum_{m=0}^{\infty} \frac{(-ik)^m}{m!}\langle x^m\rangle = \exp\left[\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}\langle x^n\rangle_c\right] = \prod_n \sum_p \left[\frac{(-ik)^{np}}{p!}\left(\frac{\langle x^n\rangle_c}{n!}\right)^p\right]$

thus $\langle x^m\rangle = \sum_{\{p\}} m! \prod_n \frac{1}{p!(n!)^p}\langle x^n\rangle_c^p$ w/ $\sum np = m$

Classical example

- Gaussian distribution. $p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\lambda)^2}{2\sigma^2}\right)$

$\tilde{p}(k) = \exp\left[-ik\lambda - \frac{k^2\sigma^2}{2}\right]$ after completing square

$\ln\tilde{p}(k) = -ik\lambda - \frac{k^2\sigma^2}{2}$

thus. $\langle x\rangle_c = \lambda \quad \langle x^2\rangle_c = \sigma^2, \quad \langle x^n\rangle_c = 0$ for $n \geq 3$

$\Rightarrow \quad \langle x\rangle = \lambda, \quad \langle x^2\rangle = \sigma^2+\lambda^2, \quad \langle x^3\rangle = 3\sigma^2\lambda + \lambda^3$

- Poisson distribution $\overset{0}{|}\overset{\leftrightarrow dt}{+++}\overset{T}{|}\longrightarrow t$

recall binomial distrib. first b/f poisson

Prob of single event is $\propto dt$: $p = \alpha dt$

" " No " " " $1-\alpha dt$ "

For consecutive $N = \frac{T}{dt}$ events, we have binomial distribution,

thus the characteristic func, $\tilde{p}(k) = (pe^{-ik}+q)^N = \lim_{dt\to 0}\left[1+\alpha dt(e^{-ik}-1)\right]^{T/dt} = \exp\left[\alpha T(e^{-ik}-1)\right]$

$$p(x) = \int_{-\infty}^{\infty}\frac{dk}{2\pi}\tilde{p}(k)e^{ikx} = \sum_{M=0}^{\infty}e^{-\alpha T}\frac{(\alpha T)^M}{M!}\delta(x-M)$$

thus $P_\lambda(M) = e^{-\lambda}\frac{\lambda^M}{M!}$ w/ $\lambda \equiv \alpha T$ is the mean!

its cumulant: $\ln\tilde{P}(k) = \alpha T(e^{-ik}-1) = \alpha T\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}$

so $\langle M^n\rangle_c = \alpha T, \quad \langle M\rangle = M_c = \alpha T = \lambda, \quad \langle M^2\rangle = \langle M^2\rangle_c + \langle M\rangle_c^2 = \alpha T + (\alpha T)^2$

- Binomial distribution. $P_A + P_B = 1$, $N = N_A + N_B$, For N events $(P_A+P_B)^N = 1$, $P_N(N_A) = \binom{N}{N_A}P_A^{N_A}P_B^{N-N_A}$

in characteristic func.

$$\tilde{P}_N(k) = \langle e^{-ikN_A}\rangle = \sum_{N_A=0}^{N}\frac{N!}{N_A!(N-N_A)!}P_A^{N_A}P_B^{N-N_A}e^{-ikN_A} = (P_Ae^{-ik}+P_B)^N$$

thus $\ln\tilde{P}_N(k) = N\ln(P_Ae^{-ik}+P_B) = N\ln(1-P_A+P_Ae^{-ik})$

here we can use expansion of $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$ & $e^x = 1 + x + \frac{x^2}{2!} + \cdots$

to extract $\langle N_A^l\rangle_c$.

Instead we consider following: Claim $\boxed{\langle N_A^l\rangle = P_A}$

$\langle N_A\rangle = P_A\frac{\partial}{\partial P_A}(P_A+P_B)^N = (N_A+N_B)P_A = P_A$ since $N_A = 0$ and 1 only

$\langle N_A^2\rangle = \left(P_A\frac{\partial}{\partial P_A}\right)\left(P_A\frac{\partial}{\partial P_A}\right)(P_A+P_B)^N = NP_A(\cancel{P_A+P_B})^{N-1} + N(\cancel{N-1})P_A^2(P_A+P_B)^{N-2} = P_A$

$\vdots$

By induction $\langle N_A^l\rangle = P_A$

Since $\ln\tilde{P}_N(k) = N\ln\tilde{P}_1(k)$ s.t.

$$\langle N_A\rangle_c = N\langle N_A\rangle = NP_A$$

$$\langle N_A^2\rangle_c = N(\langle N_A^2\rangle - \langle N_A\rangle^2) = N(P_A - P_B^2) = NP_AP_B$$

$\vdots$

Multivariables

$S_x = \{-\infty < x_1, \dots, x_N < \infty\}$, joint PDF $P_{\vec{x}}(S) = \int d^N\vec{x}\, p(\vec{x}) = 1$

for N independent random variable, $p(\vec{x}) = \prod_{i=1}^{N} p_i(x_i)$

Recall joint probability $P(A\cap B) = P(A|B)P(B) = P(B|A)P(A)$

Bayes thm $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ if A, B indept, then $P(A|B) = P(A)$ etc s.t. $P(A\cap B) = P(A)P(B)$

so for particles w/ vel. $\vec{v}$ at locations $\vec{x}$

Conditional PDF: $p(\vec{v}|\vec{x}) = \frac{p(\vec{x},\vec{v})}{p(\vec{x})}$ — joint PDF

the unconditional PDF: $p(\vec{x}) = \int d^3\vec{v}\, p(\vec{x},\vec{v})$

in general, uncond. PDF $p(x_1, \dots, x_m) = \int \prod_{i=m+1}^{N} dx_i\, p(x_1, \dots, x_N)$

cond. PDF $p(x_1, \dots, x_m | x_{m+1}, \dots, x_N) = \frac{p(x_1, \dots, x_N)}{p(x_{m+1}, \dots, x_N)}$

Now

$\langle F(\vec{x})\rangle = \int d^N\vec{x}\, p(\vec{x})\, F(\vec{x})$

characteristic func. $\tilde{p}(\vec{k}) = \left\langle \exp\left(-i\sum_{j=1}^{N} k_j x_j\right)\right\rangle$

joint moment: note $\langle x\rangle = \int x\, p(x)\, dx = \frac{\partial}{\partial(-ik)} \int dx\, p(x) e^{-ikx}\Big|_{k=0}$

thus $\langle x_1^{n_1} x_2^{n_2} \cdots x_N^{n_N}\rangle = \left[\frac{\partial}{\partial(-ik_1)}\right]^{n_1} \cdots \left[\frac{\partial}{\partial(-ik_N)}\right]^{n_N} \tilde{p}(\vec{k}=\vec{0})$

joint cumulant:

$\langle x_1^{n_1} x_2^{n_2} \cdots x_N^{n_N}\rangle_c = \left[\frac{\partial}{\partial(-ik_1)}\right]^{n_1} \cdots \left[\frac{\partial}{\partial(-ik_N)}\right]^{n_N} \ln \tilde{p}(\vec{k}=\vec{0})$

If independent.

$\int dx_1\, p(x_1) \int dx_2\, p(x_2)$

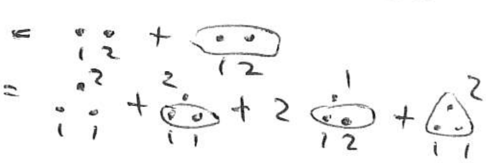
$\tilde{p}(\vec{k}) = \tilde{p}(k_1)\tilde{p}(k_2)$

$\ln \tilde{p}(\vec{k}) = \ln \tilde{p}(k_1) + \ln \tilde{p}(k_2)$

$= \sum_m \frac{(-ik_1)^m}{m!}\langle x_1^m\rangle_c + \sum_n \frac{(-ik_2)^n}{n!}\langle x_2^n\rangle_c$

thus $\langle x_1 x_2\rangle_c = \frac{\partial}{\partial(-ik_1)}\frac{\partial}{\partial(-ik_2)} \ln \tilde{p}(\vec{k})$ vanished!

Graphical rep. $\langle x_1 x_2\rangle =$; $\langle x_1^2 x_2\rangle =$



$\langle x_1 x_2\rangle = \langle x_1\rangle_c \langle x_2\rangle_c + \langle x_1 x_2\rangle_c$

$\langle x_1^2 x_2\rangle = \langle x_1\rangle_c^2 \langle x_2\rangle_c + \langle x_1^2\rangle_c \langle x_2\rangle_c + 2\langle x_1 x_2\rangle_c \langle x_1\rangle_c + \langle x_1^2 x_2\rangle_c$

note the connected correlation $\boxed{\langle x_\alpha x_\beta\rangle_c = 0 \text{ if } x_\alpha, x_\beta \text{ indept.}}$

- joint Gaussian dist. $p(\vec{x}) = \frac{1}{\sqrt{(2\pi)^N \det[C]}} \exp\left[-\frac{1}{2}\sum_{mn} (C^{-1})_{mn}(x_m - \lambda_m)(x_n - \lambda_n)\right]$ C sym.
- apply change of variable $y_j = x_j - \lambda_j$
- apply unitary diagonalization $C = A^{+}\Lambda A$ $C^{+} = A^{+}\Lambda^{-1}A$ s.t. $C^{+}C = I$ $\det C = \det \Lambda$

$\Rightarrow x^T C x = (Ax)^{+} \Lambda A x$ by completing the square, this gives normalized factor $N \sim \frac{1}{\sqrt{(2\pi)^N \det C}}$

- Also gives: $\tilde{p}(\vec{k}) = \exp\left[-ik_m\lambda_m - \frac{1}{2} C_{mn} k_m k_n\right]$
- thus we have $\langle x_m\rangle_c = \lambda_m$, $\langle x_m x_n\rangle_c = C_{mn}$

Wick's thm

- special case $\lambda_m = 0$, we have $\langle x_a x_b x_c x_d\rangle = C_{ab}C_{cd} + C_{ac}C_{bd} + C_{ad}C_{bc}$.

Central Limit Thm

Given Sum of Random variable: $X = \sum x_i$ and joint PDF $p(\vec{x})$

Then PDF of X is:

$$P_X(x) = \int d^N x \, p(\vec{x}) \, \delta(x - \sum_{i=1}^{N} x_i)$$

characteristic func. $\tilde{P}_X(k) = \langle \exp(-ikX) \rangle = \tilde{p}(k_1 = k_2 = \cdots = k_N = k)$

Cumulants $\ln \tilde{p}(k_1 = \cdots = k_N = k) = -ik \sum_{i_1=1}^{N} \langle x_{i_1} \rangle_c + \frac{(-ik)^2}{2} \sum_{i_1, i_2}^{N} \langle x_{i_1} x_{i_2} \rangle_c + \cdots$

define $\langle X \rangle_c = \sum_{i=1}^{N} \langle x_i \rangle_c$, $\langle X^2 \rangle_c = \sum_{i,j}^{N} \langle x_i x_j \rangle_c$, ...

- if independent rand. variable $p(\vec{x}) = \prod p_i(x_i) \Rightarrow \tilde{P}_X(k) = \prod \tilde{p}_i(k) \Rightarrow n$ cumulant $\langle X^n \rangle_c = \sum_{i=1}^{N} \langle x_i^n \rangle_c$
- If $p_i(x_i) = p(x_i) \Rightarrow \langle X^n \rangle_c = N \langle x^n \rangle_c \Rightarrow$ mean $\propto N$, $\sigma \sim \sqrt{N} \Rightarrow$ binomial distribution
- Now if define rand. var $y = \frac{X - N\langle x \rangle_c}{\sqrt{N}}$ which $\langle y \rangle = 0$ $\langle y^n \rangle_c \propto N^{1 - n/2}$

as $N \to \infty$ only first two cumulants survive and PDF of y becomes gaussian.

$$\lim_{N \to \infty} p\left(y = \frac{\sum_{i=1}^{N} x_i - N\langle x \rangle_c}{\sqrt{N}}\right) = \frac{1}{\sqrt{2\pi \langle x^2 \rangle_c}} \exp\left(-\frac{y^2}{2\langle x^2 \rangle_c}\right)$$

Punch line: (Central limit thm)

Convergence of PDF for sum over many rand. variables to Gaussian distribution!

Remark: independent of rand. variable is not a necessary cond. but w/ only $\sum_{i_1, i_2, \cdots i_m}^{N} \langle x_{i_1} \cdots x_{i_m} \rangle_c \ll O(N^{m/2})$

Information entropy

Event outcome $S = \{X_i\}$ $i=1,\dots,M$ corresponding to prob. $\{p(i)\}$

A message contains characters of N indep. outcomes of random variable in binary: $M = 2^{\ln_2 M}$

thus of total $N \ln_2 M$ bits.

In large N, $\{N_i\}$ occurance of $\{X_i\}_{i=1}^{M}$ is expected to be $\{N_i \approx N p_i\}$ fraction of total N

~~then~~ number of $\{N_i\}$ occurances of $\{X_i\}$:

$$g = \frac{N!}{\prod_{i=1}^{M} N_i!} < M^N$$

$$\ln_2 g = \ln N! - \sum_{i=1}^{M} \ln_2 (Np_i)! \approx -N \sum_{i=1}^{M} p_i \ln_2 p_i \quad \text{as } N \to \infty$$

one can also approach this w/

$$1 = \left(\sum_i p_i\right)^N = \sum_{\{N_i\}} N! \prod_{i=1}^{M} \frac{p_i^{N_i}}{N_i!} \approx \prod_{i=1}^{M} p_i^{Np_i} \sum_{\{N_i\}} \frac{N!}{N_i!} \overset{?}{\approx} g \prod_{i=1}^{M} p_i^{Np_i}$$

Shannon's thm. min. # of bits needs to ensure % error in N trials $\to 0$ as $N \to \infty$ is $\ln_2 g < \ln_2 N$

thus difference per trial attributed to info. content distribution:

$$I[\{p_i\}] = \ln_2 M + \sum_{i=1}^{M} p_i \ln_2 p_i$$

entropy:
$$S \equiv -\sum_{i=1}^{M} p_i \ln p_i = -\langle \ln p_i \rangle$$

$$0 \le S \le \ln M$$

if $p(i) = \delta_{ij}$; if uniform distribution $p_i = \frac{1}{M}$

Estimate of probability under unbiased estimate

i.e. all outcomes are equally likely (max. distribution).
if additional info is available the unbiased estimate can be obtained by maximizing entropy under constraints. i.e. $\langle F(x) \rangle = f$.

then
$$S = -\sum_i p_i \ln p_i - \alpha\left(\sum p_i - 1\right) - \beta\left(\sum p_i F(x_i) - f\right)$$

after maximization,
$$p_i \propto e^{-\beta F(x_i)}\,!$$