Bernoulli : 0/1 . $Pr(X=1) = p$, $Pr(X=0) = 1-p$

or, $f(x|p) = p^x (1-p)^{1-x}$

Binomial : Sum of n Bernoulli variable.

$$f(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

Hypergeometric: prob of getting x red balls when drawing n balls from a box with A red and B blue w/o replacement

$$f(x|A,B,n) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}$$

Note, if $A, B >> n$, $f(x|A,B,n) \simeq$ binomial with $p = \frac{A}{A+B}$

Poisson: number of occurrences for random arrival in unit time.

$$f(x) = \underbrace{e^{-\lambda}}_{\text{normalization term}} \frac{\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

Mean $= \lambda$, Var $= \lambda$

Interpretation: average occurrence in unit time is λ.

MGF: $e^{\lambda(e^t - 1)}$

Sum of Poisson:

$X_1 \cdots X_k$ Poisson RV w/ $\lambda_1, \cdots \lambda_k$. Let $Y = X_1 + \cdots + X_k$

Then Y is Poisson w/ $\lambda = \lambda_1 + \cdots + \lambda_k$

(From intepretation, or MGF)

Closeness of Poisson and Binomial

If n large, p small, $np \approx \lambda$, then

Binomial$(n, p) \approx$ Poisson(λ)

Interpretation: divide unit time into n small intervals,

so small that in every interval the event

happens at most 1 time.

Negative Binomial: number of failures until a fixed number of success is observed.

Binomial coeff : p # success to observe : r

failure observed : x

$$f(x|r,p) = \binom{r+x-1}{x} p^r (1-p)^x \quad (x = 0, 1, 2 \cdots)$$

Geometric: negative binomial with $r=1$

$$f(x|1,p) = p(1-p)^x \quad (x = 0, 1, 2 \cdots)$$

Negative binomial w/ r, p is sum of r geometric w/ p

Normal $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Importance: 1. Math convenience, many posterior, marginal, conditional etc will have simple form.

2. Common in experiments. (Backed by CLT?)

Gaussian Integral $\int_{-\infty}^{\infty} e^{-x^2}$

Let result $= I$

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy$$

$$= \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx\, dy$$

Change coordinate to polar (r, θ).

$$I^2 = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r\, dr\, d\theta$$

$$= \int_0^{2\pi} d\theta \cdot \int_0^{\infty} e^{-r^2} r\, dr$$

$$= 2\pi \cdot 1 = 2\pi$$

$$I = \sqrt{2\pi}$$

φ and Φ: φ: pdf of $N(0,1)$

Φ: cdf of $N(0,1)$

cdf of $N(\mu, \sigma)$: $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Sum of normal:

$X_1 \cdots X_k$ normal with $\mu_{1 \dots k}$, $\sigma_{1 \dots k}$, $Y = X_1 + \dots + X_k$

Then Y normal with $\mu_1 + \dots + \mu_k$, $\sigma_1^2 + \dots + \sigma_k^2$

Log-normal: $\log(X)$ is normal

Gamma: popular distr of positive RV.
exponential distr is a subfamily of Gamma,
which is time between successive occurrences in Poisson.

Gamma function: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$

- For $\alpha > 1$, $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$
- $\Gamma(n) = (n-1)!$
- For $\alpha, \beta > 0$, $\int_0^\infty x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$

Gamma distribution:
For $\alpha, \beta > 0$, it's continuous distr

$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} e^{-\beta x} \quad (x > 0)$$

$\mathbb{E}(X) = \alpha/\beta$. $Var(X) = \alpha/\beta^2$

Sum of Gamma:
$X_1 \cdots X_k$ Gamma RV w/ $\alpha_1 \cdots \alpha_k$, $\beta \cdots \beta$, $Y = X_1 + \cdots + X_k$
Then Y is Gamma w/ $\alpha_1 + \cdots + \alpha_k$, β

Exponential: Gamma with $\alpha = 1$

$$f(x|\beta) = \beta e^{-\beta x} \quad (x > 0)$$

Memoryless property:

X exponential, $t > 0$, $h > 0$,

$$\Pr(X \geq t+h \mid X \geq t) = \Pr(X \geq h)$$

Min of exponential RVs

$X_1 \cdots X_n$ exponential w/ β, $Y = \min\{X_i\}$,

Then Y is exponential w/ $n\beta$.

K-th smallest of exponential RVs:

$X_1 \cdots X_n$ exponential w/ β, $Z_1 \cdots Z_n$ are X_is sorted.

For $k = 2 \cdots n$, let $Y_k = Z_k - Z_{k-1}$.

Then Y_k is exponential w/ $(n+1-k)\beta$

Relation to Poisson processes

For Poisson process w/ mean β, time between arrivals are i.i.d w/ exponential distr w/ β.

Time until k-th arrival is Gamma distributed w/ k and β.

Beta: popular family of distr for RV in $[0,1]$. e.g., success rate of Bernoulli experiment

Beta function

For $\alpha, \beta > 0$,

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

- Beta function is finite for all $\alpha, \beta > 0$
- For $\alpha, \beta > 0$, $B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

Beta distr

$\alpha, \beta > 0$,

$$f(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad (0 < x < 1)$$

- If $P \sim \text{Beta}(\alpha,\beta)$, $\Pr(X|P=p)$ is $\text{Binomial}(n,p)$, then $\Pr(P|X=x)$ is $\text{Beta}(\alpha+x, \beta+n-x)$

  Beta is a conjugate prior in this case.