Probability Cheat Sheet

Distributions

Unifrom Distribution

`notation`	$U[a,b]$
`cdf`	$\frac{x-a}{b-a}$ for $x \in [a,b]$
`pdf`	$\frac{1}{b-a}$ for $x \in [a,b]$
`expectation`	$\frac{1}{2}(a+b)$
`variance`	$\frac{1}{12}(b-a)^2$
`mgf`	$\frac{e^{tb}-e^{ta}}{t(b-a)}$

`story:` all intervals of the same length on the distribution's support are equally probable.

Gamma Distribution

`notation`	$Gamma(k,\theta)$
`pdf`	$\frac{\theta^k x^{k-1} e^{-\theta x}}{\Gamma(k)} \mathbb{I}_{x>0}$ $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$
`expectation`	$k\theta$
`variance`	$k\theta^2$
`mgf`	$(1-\theta t)^{-k}$ for $t < \frac{1}{\theta}$
`ind. sum`	$\sum_{i=1}^n X_i \sim Gamma\left(\sum_{i=1}^n k_i, \theta\right)$

`story:` the sum of k independent exponentially distributed random variables, each of which has a mean of θ (which is equivalent to a rate parameter of θ^{-1}).

Geometric Distribution

`notation`	$G(p)$
`cdf`	$1-(1-p)^k$ for $k \in \mathbb{N}$
`pmf`	$(1-p)^{k-1} p$ for $k \in \mathbb{N}$
`expectation`	$\frac{1}{p}$
`variance`	$\frac{1-p}{p^2}$
`mgf`	$\frac{pe^t}{1-(1-p)e^t}$

`story:` the number X of Bernoulli trials needed to get one success. Memoryless.

Poisson Distribution

`notation`	$Poisson(\lambda)$
`cdf`	$e^{-\lambda} \sum_{i=0}^k \frac{\lambda^i}{i!}$
`pmf`	$\frac{\lambda^k}{k!} \cdot e^{-\lambda}$ for $k \in \mathbb{N}$
`expectation`	λ
`variance`	λ
`mgf`	$\exp\left(\lambda\left(e^t-1\right)\right)$
`ind. sum`	$\sum_{i=1}^n X_i \sim Poisson\left(\sum_{i=1}^n \lambda_i\right)$

`story:` the probability of a number of events occurring in a fixed period of time if these events occur with a known average rate and independently of the time since the last event.

Normal Distribution

`notation`	$N\left(\mu,\sigma^2\right)$
`pdf`	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/\left(2\sigma^2\right)}$
`expectation`	μ
`variance`	σ^2
`mgf`	$\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$
`ind. sum`	$\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$

`story:` describes data that cluster around the mean.

Standard Normal Distribution

`notation`	$N(0,1)$
`cdf`	$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$
`pdf`	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
`expectation`	$\frac{1}{\lambda}$
`variance`	$\frac{1}{\lambda^2}$
`mgf`	$\exp\left(\frac{t^2}{2}\right)$

`story:` normal distribution with $\mu = 0$ and $\sigma = 1$.

Exponential Distribution

`notation`	$exp(\lambda)$
`cdf`	$1-e^{-\lambda x}$ for $x \geq 0$
`pdf`	$\lambda e^{-\lambda x}$ for $x \geq 0$
`expectation`	$\frac{1}{\lambda}$
`variance`	$\frac{1}{\lambda^2}$
`mgf`	$\frac{\lambda}{\lambda - t}$
`ind. sum`	$\sum_{i=1}^k X_i \sim Gamma(k,\lambda)$
`minimum`	$\sim exp\left(\sum_{i=1}^k \lambda_i\right)$

`story:` the amount of time until some specific event occurs, starting from now, being memoryless.

Binomial Distribution

`notation`	$Bin(n,p)$
`cdf`	$\sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}$
`pmf`	$\binom{n}{i} p^i (1-p)^{n-i}$
`expectation`	np
`variance`	$np(1-p)$
`mgf`	$\left(1-p+pe^t\right)^n$

`story:` the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p.

Basics

Comulative Distribution Function

$F_X(x) = \mathbb{P}(X \leq x)$

Probability Density Function

$F_X(x) = \int_{-\infty}^{\infty} f_X(t)\, dt$

$\int_{-\infty}^{\infty} f_X(t)\, dt = 1$

$f_X(x) = \frac{d}{dx} F_X(x)$

Quantile Function

The function $X^* : [0,1] \to \mathbb{R}$ for which for any $p \in [0,1]$, $F_X\left(X^*(p)^-\right) \leq p \leq F_X\left(X^*(p)\right)$

$F_{X^*} = F_X$

$\mathbb{E}(X^*) = \mathbb{E}(X)$

Expectation

$\mathbb{E}(X) = \int_0^1 X^*(p) dp$

$\mathbb{E}(X) = \int_{-\infty}^0 F_X(t)\, dt + \int_0^\infty (1-F_X(t))\, dt$

$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X x dx$

$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X x dx$

$\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$

Variance

$\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2$

$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$

$\mathrm{Var}(aX+b) = a^2 \mathrm{Var}(X)$

Standard Deviation

$\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Covariance

$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

$\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mathbb{E}(x))(Y-\mathbb{E}(Y)))$

$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$

Correlation Coefficient

$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X, \sigma_Y}$

Moment Generating Function

$M_X(t) = \mathbb{E}\left(e^{tX}\right)$

$\mathbb{E}(X^n) = M_X^{(n)}(0)$

$M_{aX+b}(t) = e^{tb} M_{aX}(t)$

Joint Distribution

$\mathbb{P}_{X,Y}(B) = \mathbb{P}((X,Y) \in B)$
$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y)$

Joint Density

$$\mathbb{P}_{X,Y}(B) = \iint_B f_{X,Y}(s,t)\,dsdt$$

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,dtds$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,dsdt = 1$$

Marginal Distributions

$\mathbb{P}_X(B) = \mathbb{P}_{X,Y}(B \times \mathbb{R})$
$\mathbb{P}_Y(B) = \mathbb{P}_{X,Y}(\mathbb{R} \times Y)$

$$F_X(a) = \int_{-\infty}^{a}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,dtds$$

$$F_Y(b) = \int_{-\infty}^{b}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,dsdt$$

Marginal Densities

$$f_X(s) = \int_{-\infty}^{\infty} f_{X,Y}(s,t)dt$$

$$f_Y(t) = \int_{-\infty}^{\infty} f_{X,Y}(s,t)ds$$

Joint Expectation

$$\mathbb{E}(\varphi(X,Y)) = \iint_{\mathbb{R}^2} \varphi(x,y)\, f_{X,Y}(x,y)\,dxdy$$

Independent r.v.

$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\,\mathbb{P}(Y \leq y)$
$F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$
$f_{X,Y}(s,t) = f_X(s)\,f_Y(t)$
$\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$
$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
Independent events:
$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$

Conditional Probability

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

`bayes` $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)}$

Conditional Density

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{X|Y=n}(x) = \frac{f_X(x)\,\mathbb{P}(Y=n \mid X=x)}{\mathbb{P}(Y=n)}$$

$$F_{X|Y=y} = \int_{-\infty}^{x} f_{X|Y=y}(t)\,dt$$

Conditional Expectation

$$\mathbb{E}(X \mid Y=y) = \int_{-\infty}^{\infty} x f_{X|Y=y}(x)\,dx$$

$\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(X)$
$\mathbb{P}(Y=n) = \mathbb{E}(\mathbb{I}_{Y=n}) = \mathbb{E}(\mathbb{E}(\mathbb{I}_{Y=n} \mid X))$

Sequences and Limits

$$\limsup A_n = \{A_n \text{ i.o.}\} = \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n$$

$$\liminf A_n = \{A_n \text{ eventually}\} = \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} A_n$$

$\liminf A_n \subseteq \limsup A_n$
$(\limsup A_n)^c = \liminf A_n^c$
$(\liminf A_n)^c = \limsup A_n^c$

$$\mathbb{P}(\limsup A_n) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{n=m}^{\infty} A_n\right)$$

$$\mathbb{P}(\liminf A_n) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{n=m}^{\infty} A_n\right)$$

Borel-Cantelli Lemma

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(\limsup A_n) = 0$$

And if A_n are independent:

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}(\limsup A_n) = 1$$

Convergence

Convergence in Probability

`notation` $X_n \xrightarrow{p} X$

`meaning` $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$

Convergence in Distribution

`notation` $X_n \xrightarrow{D} X$

`meaning` $\lim_{n\to\infty} F_n(x) = F(x)$

Almost Sure Convergence

`notation` $X_n \xrightarrow{a.s.} X$

`meaning` $\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1$

Criteria for a.s. Convergence

- $\forall\varepsilon\exists N\forall n > N : \mathbb{P}(|X_n - X| < \varepsilon) > 1 - \varepsilon$
- $\forall\varepsilon \mathbb{P}(\limsup(|X_n - X| > \varepsilon)) = 0$
- $\forall\varepsilon \sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$ (by B.C.)

Convergence in L_p

`notation` $X_n \xrightarrow{L_p} X$

`meaning` $\lim_{n\to\infty} \mathbb{E}(|X_n - X|^p) = 0$

Relationships

$$\begin{array}{ccccc} \xrightarrow{L_q} & \underset{q>p\geq 1}{\Rightarrow} & \xrightarrow{L_p} & & \\ & & \Downarrow & & \\ \xrightarrow{a.s.} & \Rightarrow & \xrightarrow{p} & \Rightarrow & \xrightarrow{D} \end{array}$$

If $X_n \xrightarrow{D} c$ then $X_n \xrightarrow{p} c$
If $X_n \xrightarrow{p} X$ then there exists a subsequence n_k s.t. $X_{n_k} \xrightarrow{a.s.} X$

Laws of Large Numbers

If X_i are i.i.d. r.v.,

`weak law` $\overline{X_n} \xrightarrow{p} \mathbb{E}(X_1)$

`strong law` $\overline{X_n} \xrightarrow{a.s.} \mathbb{E}(X_1)$

Central Limit Theorem

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1)$$

If $t_n \to t$, then

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq t_n\right) \to \Phi(t)$$

Inequalities

Markov's inequality

$$\mathbb{P}(|X| \geq t) \leq \frac{\mathbb{E}(|X|)}{t}$$

Chebyshev's inequality

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}$$

Chernoff's inequality

Let $X \sim Bin(n,p)$; then:
$\mathbb{P}(X - \mathbb{E}(X) > t\sigma(X)) < e^{-t^2/2}$
Simpler result; for every X:
$\mathbb{P}(X \geq a) \leq M_X(t)\,e^{-ta}$

Jensen's inequality

for φ a convex function, $\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$

Miscellaneous

$$\mathbb{E}(Y) < \infty \iff \sum_{n=0}^{\infty} \mathbb{P}(Y > n) < \infty \ (Y \geq 0)$$

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} \mathbb{P}(X > n) \ (X \in \mathbb{N})$$

$X \sim U(0,1) \iff -\ln X \sim exp(1)$

Convolution

For ind. X, Y, $Z = X + Y$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(s)\,f_Y(z-s)\,ds$$

Kolmogorov's 0-1 Law

If A is in the tail σ-algebra $\mathcal{F}^t$, then $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$

Ugly Stuff

cdf of Gamma distribution:

$$\int_0^t \frac{\theta^k x^{k-1} e^{-\theta k}}{(k-1)!}dx$$

This cheatsheet was made by Peleg Michaeli in November 2010, using LaTeX.
version: 1.01
comments: peleg.michaeli@math.tau.ac.il