Probability Cheat Sheet

Distributions

Unifrom Distribution

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

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

Gamma Distribution

`notation`	$Gamma\left(k,\theta\right)$
`pdf`	$\frac{\theta^{k}x^{k-1}e^{-\theta x}}{\Gamma\left(k\right)}\mathbb{I}_{x>0}$ $\Gamma\left(k\right)=\int_{0}^{\infty}x^{k-1}e^{-x}dx$
`expectation`	$k\theta$
`variance`	$k\theta^{2}$
`mgf`	$\left(1-\theta t\right)^{-k}$ 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`	$G\left(p\right)$
`cdf`	$1-\left(1-p\right)^{k}$ for $k\in\mathbb{N}$
`pmf`	$\left(1-p\right)^{k-1}p$ for $k\in\mathbb{N}$
`expectation`	$\frac{1}{p}$
`variance`	$\frac{1-p}{p^{2}}$
`mgf`	$\frac{pe^{t}}{1-\left(1-p\right)e^{t}}$

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

Poisson Distribution

`notation`	$Poisson\left(\lambda\right)$
`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`	$\frac{1}{\sqrt{2\pi\sigma^{2}}}e^{-\left(x-\mu\right)^{2}/\left(2\sigma^{2}\right)}$
`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`	$N\left(0,1\right)$
`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`	$exp\left(\lambda\right)$
`cdf`	$1-e^{-\lambda x}$ for $x\geq0$
`pdf`	$\lambda e^{-\lambda x}$ for $x\geq0$
`expectation`	$\frac{1}{\lambda}$
`variance`	$\frac{1}{\lambda^{2}}$
`mgf`	$\frac{\lambda}{\lambda-t}$
`ind. sum`	$\sum_{i=1}^{k}X_{i}\sim Gamma\left(k,\lambda\right)$
`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`	$\sum_{i=0}^{k}\binom{n}{i}p^{i}\left(1-p\right)^{n-i}$
`pmf`	$\binom{n}{i}p^{i}\left(1-p\right)^{n-i}$
`expectation`	np
`variance`	$np\left(1-p\right)$
`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

$F_{X}\left(x\right)=\mathbb{P}\left(X\leq x\right)$

Probability Density Function

$F_{X}\left(x\right)=\int_{-\infty}^{\infty}f_{X}\left(t\right)dt$

$\int_{-\infty}^{\infty}f_{X}\left(t\right)dt=1$

$f_{X}\left(x\right)=\frac{d}{dx}F_{X}\left(x\right)$

Quantile Function

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

$F_{X^{*}}=F_{X}$

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

Expectation

$\mathbb{E}\left(X\right)=\int_{0}^{1}X^{*}(p)dp$

$\mathbb{E}\left(X\right)=\int_{-\infty}^{0}F_{X}\left(t\right)dt+\int_{0}^{\infty}\left(1-F_{X}\left(t\right)\right)dt$

$\mathbb{E}\left(X\right)=\int_{-\infty}^{\infty}xf_{X}xdx$

$\mathbb{E}\left(g\left(X\right)\right)=\int_{-\infty}^{\infty}g\left(x\right)f_{X}xdx$

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

Variance

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

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

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

Standard Deviation

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

Covariance

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

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

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

Correlation Coefficient

$\rho_{X,Y}=\frac{\mathrm{Cov}\left(X,Y\right)}{\sigma_{X},\sigma_{Y}}$

Moment Generating Function

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

$\mathbb{E}\left(X^{n}\right)=M_{X}^{(n)}\left(0\right)$

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

Joint Distribution

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

Joint Density

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

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

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

Marginal Distributions

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

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

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

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

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

Independent r.v.

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

Conditional Probability

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

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

Conditional Density

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

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

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

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

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

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

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

If $t_n \to t$, then

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

Inequalities

Markov's inequality

$\mathbb{P}(|X| \ge t) \le \dfrac{\mathbb{E}(|X|)}{t}$

Chebyshev's inequality

$\mathbb{P}(|X - \mathbb{E}(X)| \ge \varepsilon) \le \dfrac{\mathrm{Var}(X)}{\varepsilon^2}$

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:
$\mathbb{P}(X \ge a) \le M_X(t)\, e^{-ta}$

Jensen's inequality

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

Miscellaneous

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

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

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

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:

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

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