Formula Sheet

1. Basic sums: $\sum_{i=1}^n i = \frac{1}{2}n(n+1)$, $\sum_{i=1}^n i^2 = \frac{1}{6}n(n+1)(2n+1)$, $\sum_{i=1}^n i^3 = \frac{1}{4}n^2(n+1)^2$, $n \geq 1$.
2. Binomial coefficients: $\binom{n}{k} := \frac{n!}{(n-k)!k!} = \binom{n}{n-k}$, $n \geq k \geq 0$, $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$, $n, k \geq 1$.
3. Sums of them: $\sum_{i=0}^n \binom{k+i-1}{i} = \binom{n+k}{k}$, $k, n \geq 0$, $\sum_{i=0}^n \binom{l}{i}\binom{m-l}{n-i} = \binom{m}{n}$, $m \geq 0$, $n, l = 0, \ldots, m$.
4. Binomial theorem: $\sum_{i=0}^n \binom{n}{i} a^i b^{n-i} = (a+b)^n$, $n \geq 0$, $a, b \in \mathbb{R}$.
5. Sum of a geometric progression: $\sum_{i=0}^n a^i = \frac{1-a^{n+1}}{1-a}$ if $a \neq 1$, and $\sum_{i=0}^n a^i = n+1$ if $a = 1$.
6. Derivative: $f'(a) := \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}$.
7. Differentiation: $\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx}$, $\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx}$, $\frac{d}{dx}(f \circ g) = \frac{df}{dg}\frac{dg}{dx}$ $\left((f \circ g)(x) := f(g(x))\right)$.
8. Integration by parts: if $g(x) = G'(x)$, then $\int_a^b f(x)g(x)\,dx = f(b)G(b) - f(a)G(a) - \int_a^b f'(x)G(x)\,dx$.
9. Taylor's theorem: if $f, f', \ldots, f^{(n+1)}$ are defined on $[a, x]$, then (using the Lagrange remainder)

   $$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(t)}{(n+1)!}(x-a)^{n+1} \text{ for some } t \in (a, x).$$
10. Triangle inequalities: $||x| - |y|| \leq |x+y| \leq |x| + |y|$.
11. Logarithm: $\ln(x) = \int_1^x \frac{1}{t}\,dt$, $x > 0$, $\ln(xy) = \ln(x) + \ln(y)$, $\ln(y^\alpha) = \alpha\ln(y)$, $\log_a(x) = \ln(x)/\ln(a)$.
12. Exponential: $\exp := \ln^{-1}$, $e := \exp(1)$, $e^x := \exp(x)$, $e^{x+y} = e^x e^y$, $a^x := e^{x\ln(a)}$, $a^{x+y} = a^x a^y$, $a > 0$.

    $$\frac{d}{dx}e^{ax} = ae^{ax},\ a \in \mathbb{R}, \quad \frac{d}{dx}a^{bx} = b\ln(a)a^{bx},\ a > 0,\ \sum_{i=0}^\infty \frac{a^i}{i!} = e^a,\ \lim_{n\to\infty}\left(1 + \frac{a}{n}\right)^n = e^a,\ a \in \mathbb{R}.$$
13. Geometric series: $\sum_{i=0}^\infty a^i = \frac{1}{1-a}$, $\sum_{i=1}^\infty ia^{i-1} = \left(\frac{1}{1-a}\right)^2$, $|a| < 1$, $\sum_{i=1}^\infty \frac{a^i}{i} = \ln\left(\frac{1}{1-a}\right)$, $-1 \leq a < 1$.
14. Binomial series: $\sum_{i=0}^\infty \binom{n+i-1}{i}(-1)^i a^i = \left(\frac{1}{1+a}\right)^n$, $n \geq 1$, $|a| < 1$.
15. Trigonometric functions: $\tan x = \frac{\sin x}{\cos x}$, $\sin(2x) = 2\sin x\cos x$, $\cos(2x) = 1 - 2\sin^2 x = 2\cos^2 x - 1$,
    $\sin(x + \frac{\pi}{2}) = \cos x$, $\sin(x + 2k\pi) = \sin x$, $\cos(x + 2k\pi) = \cos x$, $\tan(x + k\pi) = \tan x$,
    $\sin^2 x + \cos^2 x = 1$, $\frac{d}{dx}\sin x = \cos x$, $\frac{d}{dx}\cos x = -\sin x$, $\frac{d}{dx}\tan x = 1 + \tan^2 x$,
    $\sin(x+y) = \sin x\cos y + \cos x\sin y$, $\cos(x+y) = \cos x\cos y - \sin x\sin y$,
    $\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x\tan y}$, $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$, $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$.
16. Gamma function: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$, $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(n+1) = n!$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$,
17. Beta function: $\beta(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, $a, b > 0$.
18. Sterling's formula: $n! \sim \sqrt{2\pi}n^{n+1/2}e^{-n}$ as $n \to \infty$, $n! = \sqrt{2\pi}n^{n+1/2}e^{(-n+\frac{t}{12n})}$ for some $t \in (0, 1)$.
19. Probability. Ω is the sample space, $A, B, \cdots$ are events, $P(\cdot)$ is probability measure, $X, Y, \cdots$ are random variables (rvs), S is the range of X if X is a discrete rv, F_X denotes distribution function, f_X denotes probability density function (pdf) when X is a continuous rv, and, $\mathbb{E}(\cdot)$ is expectation.
    (a) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$, $P(A|B) := P(A\cap B)/P(B)$, $P(B|A) = P(A|B)P(B)/P(A)$.
    (b) Total probability: $P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$, where $\{B_i\}$ is a partition of Ω.
    (c) $F_X(x) := \Pr(X \leq x)$. If X is a continuous rv, $F_X(x) = \int_{-\infty}^x f_X(u)\,du$.
    (d) $\mathbb{E}(X) = \sum_{x\in S} x\Pr(X = x)$ (discrete), $\mathbb{E}(X) = \int_{-\infty}^\infty xf_X(x)\,dx$ (continuous).
    (e) $\mathbb{E}(g(X)) = \sum_{x\in S} g(x)\Pr(X = x)$ (discrete), $\mathbb{E}(g(X)) = \int_{-\infty}^\infty g(x)f_X(x)\,dx$ (continuous).
    (f) $\mathrm{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$, $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$.
    (g) $\mathrm{Cov}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
    (h) Conditional expectation: $\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$, $\mathbb{E}(Yg(Z)|Z) = g(Z)\mathbb{E}(Y|Z)$.
    (i) Probability generating function (pgf): if X is a non-negative discrete rv, $G_X(z) = \mathbb{E}(z^X)$.
    (j) Moment generating function (mgf): if $\mathbb{E}(|X|^k) < \infty$ for all k, $M_X(t) = \mathbb{E}(e^{tX})$.
    (k) Laplace-Steiltjes transform (LST): if X is a non-negative rv, $L_X(t) = \mathbb{E}(e^{-tX})$, $t \geq 0$.
    (l) Characteristic function (cf): if X is *any* rv, $\phi_X(t) = \mathbb{E}(e^{itX})$, $t \in \mathbb{R}$ (here $i = \sqrt{-1}$).

20. Discrete distributions: Here X is a discrete rv taking values in a denumerable set. The mean, variance and probability function are listed, together with the pgf $G(z) = \mathbb{E}(z^X)$, $|z| \le 1$.

Constant $\Pr(X = c) = 1$, $\mathbb{E}(X) = c$, $\mathrm{Var}(X) = 0$, $G(z) = z^c$.

Binomial ($B(n,p)$: $0 < p < 1$, $n \ge 1$) $\mathbb{E}(X) = np$, $\mathrm{Var}(X) = np(1-p)$,

$$\Pr(X = j) = \binom{n}{j} p^j (1-p)^{n-j}, \; j \in \{0, 1, \ldots, n\}, \qquad G(z) = (1 - p + pz)^n.$$

The *Bernoulli* distribution is the special case $B(1, p)$.

Poisson (*Poisson*(λ): $\lambda > 0$) $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$,

$$\Pr(X = j) = e^{-\lambda} \frac{\lambda^j}{j!}, \; j \in \{0, 1, \ldots\}, \qquad G(z) = e^{-\lambda(1-z)}.$$

Geometric ($0 < q < 1$) $\mathbb{E}(X) = q/(1-q)$, $\mathrm{Var}(X) = q/(1-q)^2$,

$$\Pr(X = j) = (1-q)q^j, \; j \in \{0, 1, \ldots\}, \quad \big(\text{Note: } \Pr(X \ge j) = q^j\big) \qquad G(z) = \frac{1-q}{1-qz}.$$

Negative binomial ($0 < q < 1$, $n \ge 1$) $\mathbb{E}(X) = nq/(1-q)$, $\mathrm{Var}(X) = nq/(1-q)^2$,

$$\Pr(X = j) = \binom{n+j-1}{j}(1-q)^n q^j, \; j \in \{0, 1, \ldots\}, \qquad G(z) = \left(\frac{1-q}{1-qz}\right)^n.$$

Hypergeometric ($N \ge 0$, $0 \le n, a \le N$) $\mathbb{E}(X) = na/N$, $\mathrm{Var}(X) = na(N-n)(N-a)/(N^2(N-1))$,

$$\Pr(X = j) = \binom{a}{j}\binom{N-a}{n-j} \Big/ \binom{N}{n}, \; j \in \{\max(0, n+a-N), \ldots, \min(n, a)\}, \qquad G(z) = \textit{complicated.}$$

21. Continuous distributions: Here X is a continuous rv taking values in a subset of $\mathbb{R}$. The mean, variance, pdf $f : \mathbb{R} \to [0, \infty)$ and (if it can be written down explicitly) the distribution function $F : \mathbb{R} \to [0, 1]$ are listed; f takes the value 0 outside the range given, so that F takes the value 0 below that range and 1 above. The mgf $M(t) = \mathbb{E}(e^{tX})$, or cf $\phi(t) = \mathbb{E}(e^{itX})$, whichever is appropriate, is also listed. For non-negative rvs, the LST satisfies $L(t) = \mathbb{E}(e^{-tX}) = M(-t)$, $t \ge 0$.

Uniform ($U(a,b)$: $a < b$) $\mathbb{E}(X) = (a+b)/2$, $\mathrm{Var}(X) = (b-a)^2/12$,

$$f(x) = \frac{1}{b-a}, \; F(x) = \frac{x-a}{b-a}, \; a \le x \le b, \qquad M(0) = 1, \; M(t) = \frac{e^{bt} - e^{at}}{(b-a)t}, \; t \ne 0.$$

Exponential ($\exp(\lambda)$: $\lambda > 0$) $\mathbb{E}(X) = 1/\lambda$, $\mathrm{Var}(X) = 1/\lambda^2$,

$$f(x) = \lambda e^{-\lambda x}, \; F(x) = 1 - e^{-\lambda x}, \; x \ge 0, \qquad M(t) = \frac{\lambda}{\lambda - t}, \; t < \lambda.$$

Gamma ($\Gamma(\alpha, \lambda)$: $\alpha > 0$, $\lambda > 0$) $\mathbb{E}(X) = \alpha/\lambda$, $\mathrm{Var}(X) = \alpha/\lambda^2$,

$$f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, \; x \ge 0, \qquad M(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}, \; t < \lambda.$$

The *Chi-squared* distribution χ^2_n ($n \ge 1$) is $\Gamma(n/2, 1/2)$. The *Erlang* distribution is $\Gamma(n, \lambda)$, and

$$f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!}, \; F(x) = 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!}, \; x \ge 0.$$

Beta ($a > 0$, $b > 0$) $\mathbb{E}(X) = a/(a+b)$, $\mathrm{Var}(X) = ab/((a+b)^2(a+b+1))$,

$$f(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}, \; 0 \le x \le 1, \qquad M(t) = \textit{complicated.}$$

Normal (Gaussian) ($N(\mu, \sigma^2)$: $\mu \in \mathbb{R}$, $\sigma^2 > 0$) $\mathbb{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\tfrac{1}{2}(x-\mu)^2/\sigma^2\right), \; x \in \mathbb{R}, \qquad M(t) = \exp\left(\mu t + \tfrac{1}{2}\sigma^2 t^2\right), \; t \in \mathbb{R}.$$

Multivariate Normal ($N(\mu, V)$: $\mu \in \mathbb{R}^n$, V +ve-definite symmetric) $\mathbb{E}(X) = \mu$, $\mathrm{Cov}(X) = V$,

$$f(x) = \frac{1}{\sqrt{(2\pi)^n |V|}} \exp\left(-\tfrac{1}{2}(x-\mu)V^{-1}(x-\mu)^T\right), \; x \in \mathbb{R}^n, \quad M(t) = \exp\left(\mu t^T + \tfrac{1}{2} t V t^T\right), \; t \in \mathbb{R}^n.$$

Cauchy ($m \in \mathbb{R}$, $b > 0$) median$= m$ (Note that $\mathbb{E}(X)$ does not exist: $\mathbb{E}(X^+) = \mathbb{E}(X^-) = \infty$)

$$f(x) = \frac{b}{\pi(b^2 + (x-m)^2)}, \; F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x-m}{b}\right), \; x \in \mathbb{R}, \qquad \phi(t) = e^{imt - b|t|}, \; t \in \mathbb{R}.$$

Weibull ($\lambda > 0$, $\beta > 0$) $\mathbb{E}(X) = \lambda^{-1/\beta}\Gamma(1 + 1/\beta)$, $\mathrm{Var}(X) = \lambda^{-2/\beta}\{\Gamma(1 + 2/\beta) - (\Gamma(1 + 1/\beta))^2\}$,

$$f(x) = \lambda\beta x^{\beta-1} \exp\left(-\lambda x^{\beta}\right), \; F(x) = 1 - \exp\left(-\lambda x^{\beta}\right), \; x \ge 0, \qquad M(t) = \textit{complicated.}$$

Laplace ($\alpha \in \mathbb{R}$, $\beta > 0$) $\mathbb{E}(X) = \alpha$, $\mathrm{Var}(X) = 2\beta^2$,

$$f(x) = \frac{1}{2\beta} \exp\left(-|x-\alpha|/\beta\right), \; x \in \mathbb{R}, \qquad M(t) = \frac{e^{\alpha t}}{1 - \beta^2 t^2}, \; |t| < 1/\beta.$$