# Summation identities

- $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$
- $\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k$ for $|x| < 1$
- $\sum_{x=1}^{n} x = \frac{n(n+1)}{2}$
- $\sum_{x=1}^{n} x^2 = \frac{n(n+1)(2n+1)}{6}$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a+b)^n$ (binomial theorem)
- $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$
- $\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}$
- $\sum_{k=1}^{n-1} x^k = \frac{x^n-1}{x-1}$ applicable when $x \neq 1$
- $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ for $x \in (-1, 1)$
- $\sum_{k=1}^{n} \frac{1}{2^k} \to 1$
- $\int_x^{\infty} \frac{1}{\Gamma(\alpha)} z^{\alpha-1} e^{-z} dz = \sum_{y=0}^{\alpha-1} \frac{x^y e^{-x}}{y!}$ $\alpha = 1, 2, 3 \ldots$

# Integrals

- Integration by parts

$$\int_a^b u\, dv = uv|_a^b - \int_a^b v\, du$$

- Median

$$\int_{-\infty}^{m} f_X(x)\, dx = \int_m^{\infty} f_X(x)\, dx$$

- Odd functions: for $f_X$ odd and $g_X$ even:

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-\infty}^{\infty} f_X(x) \cdot g_X(x)\, dx = 0$$

- For continuous even functions such that $f(-x) = f(x)$,

$$\int_{-a}^{a} f_X(x) dx = 2 \int_0^a f_X(x) dx$$

- For continuous even functions such that $f(-x) = -f(x)$,

$$\int_{-a}^{a} f_X(x)dx = 0$$

- Integral of $\int_0^\infty x^n e^{-ax}dx$, use $u$ substitution $u = ax$

$$\frac{1}{a^{n+1}} \int_0^\infty u^n e^{-u} du = \frac{\Gamma(n+1)}{a^{n+1}}$$

- $\displaystyle\int_{-\infty}^{\infty} x^k e^{-\frac{x^2}{2}} = \begin{cases} 0 & k = 1, 3, 5, 7, \ldots \\ \sqrt{2\pi} & k = 0, 2 \\ 3\sqrt{2\pi} & k = 4 \\ 15\sqrt{2\pi} & k = 6 \end{cases}, \int_0^\infty x^k e^{-\frac{x^2}{2}} = \begin{cases} \sqrt{\frac{\pi}{2}} & k = 0 \\ 1 & k = 1 \\ \sqrt{\frac{\pi}{2}} & k = 2 \\ 2 & k = 3 \\ 3\sqrt{\frac{\pi}{2}} & k = 4 \end{cases}$
- $\int_0^\infty xe^{-ax} = \frac{1}{a^2}\ a > 0$
- $\int_0^\infty x^a e^{-bx} = b^{a-1}\Gamma(a+1)$
- $\int ln(x)dx = x\, ln(x) - x = x\,(ln(x) - 1)$
- For two independent r.v.s $X$ and $Y$, the distribution of $X + Y$ is the convolution (think of $s$ as sum of $X$ and $Y$)

$$f_{X+Y}(s) = f_X * f_Y = \int_{\mathbb{R}} f_X(x) f_Y(s-x)dx$$

  - For $Z = X - Y = X + (-Y)$, $f_Z(z) = \int_{\mathbb{R}} f_X(x) f_{-Y}(z-x)dx = \int_{\mathbb{R}} f_X(x) f_Y(x-z)dx$

## Limiting identities

- $\displaystyle\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$
- For a fixed integer $m$ $\displaystyle\lim_{n\to\infty} \frac{n!(n+1)^m}{n+m} = 1$

## Other

- Identity function
  - $E[I_{[a,\infty)}(x)] = P(X \geq a)$
  - $Z_i = I(X_i < c) \sim \text{Bern}(p)$ where $p = P(X_i < c) = F(c)$. Sum of $n$ iid indicators $I_{(x\in A)}(x)$ is distributed as $\text{Bin}(n, P(x \in A))$
- Beta function: $\text{B}(a, b) = \int_0^1 u^{a-1}(1-u)^{b-1}du = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$
  - $\text{B}(x, y) = \text{B}(x, y+1) + \text{B}(x+1, y)$
  - $\text{B}(x+1, y) = \text{B}(x, y) \cdot \frac{x}{x+y}$, $\text{B}(x, y+1) = \text{B}(x, y) \cdot \frac{y}{x+y}$
- Gamma function: $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt\ z > 0$
  - $\Gamma(n) = (n-1)!$ for positive integer $n$
  - $\Gamma(1) = 1$

  - $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
  - $\Gamma(x+1) = x\Gamma(x)$

- L'Hôpital: $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)}$
- Quotient rule: $\left[\dfrac{u(x)}{v(x)}\right]' = \dfrac{u'(x)v(x) - u(x)v'(x)}{v(x)^2}$
- Product rule: $\frac{d}{dx}(u \cdot v) = v\frac{du}{dx} + u\frac{dv}{dx}$
- Taylor Series: for a function $f(x)$
  - $f(a) + \dfrac{f'(a)}{1!}(x-a) + \dfrac{f''(a)}{2!}(x-a)^2 + \dfrac{f'''(a)}{3!}(x-a)^3 + \ldots = \sum\limits_{n=0}^{\infty} \dfrac{f^{(n)}(a)}{n!}(x-a)^n$
- Completing square: $(x+p)^2 = (x^2 + 2px + p^2)$
- $x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + x^{n-2}y^2 + \ldots + xy^{n-2} + y^{n-1})$

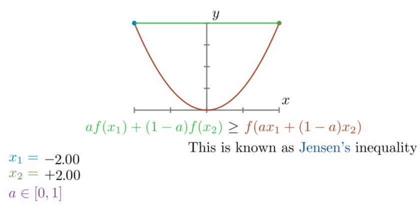
# No Brainer Quick Reference

- $Var(X) = E(X^2) - E(X)^2$
- $Var(X) = E(Var(X|Y)) + Var(E(X|Y))$
- $\sum\limits_{i=1}^{n} (x_i - \theta)^2 = n(\bar{x} - \theta)^2 + \sum\limits_{i=1}^{n} (x_i - \bar{x})^2$
- $\sum\limits_{i=1}^{n} (x_i - \theta)^2 = \sum\limits_{i=1}^{n} x_i^2 - \sum\limits_{x=1}^{n} x_i\theta + n\theta^2$
- Distributions
  - If $X \sim$ Unif$(a, b)$ then single order statistics (Casella example 5.4.5): $\frac{X_{(j)}-a}{b-a} \sim$ Beta$(j, n-j+1)$
  - First order statistics $X_{(1)}$: $f_{X_{(1)}}(x) = nf_X(x)[1 - F_X(x)]^{n-1}$
  - Last order statistics $X_{(n)}$: $f_{X_{(n)}}(x) = nf_X(x)[F_X(x)]^{n-1}$
  - Conditional distribution: $f_{Y|X}(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$
- Moment Generating Function
  - $M_X(t) = E[e^{tX}]$
  - $E[X^n] = M_x^n(0) = \frac{d^n M_X}{dt^n}(0)$
    - Discrete RV examples
      - $P_X(k) = \begin{cases} \frac{1}{3} & k = 1 \\ \frac{2}{3} & k = 2 \end{cases}$, $M_x(t) = \frac{1}{3}e^t + \frac{2}{3}e^{2t}$
      - $P_{X_n}(k) = \frac{1}{2^n}$ where $X_n = \frac{k}{2^n}$ with $k = 0, 1, \ldots, 2^n - 1$,

$$M_x(t) = E[e^{tX_n}] = \frac{1}{2^n}\sum_{k=0}^{2^n-1} e^{tk/2^n} = \frac{1}{2^n}\frac{\left(e^{t/2^n}\right)^{2^n} - 1}{e^{t/2^n} - 1}$$

$$= \frac{e^t - 1}{t}\frac{t/2^n}{e^{t/2^n} - 1} \to \frac{e^t - 1}{t} = \int_0^1 e^{tu}du = E(e^{tU})$$

  - MGF for iid samples $X_1, \ldots, X_n$
    - $M_{X_1+\ldots+X_n}(t) = [M_{X_1}(t)]^n$, $M_{\bar{X}}(t) = [M_{X_1}(t/n)]^n$
- Conditional Expectation
  - $E[E[X|Y]] = E[X]$
  - $E[X|Y] = E[X]$ if $X$ independent of $Y$
  - $E[aX + bZ|Y] = aE[X|Y] + bE[Z|Y]$
  - $E[X|Y] \geq 0$ if $x \geq 0$
  - $E[X \cdot g(Y)|Y] = g(Y) \cdot E[X|Y]$ and $E[g(Y)|Y] = g(Y)$
  - $E[X|Y, g(Y)] = E[X|Y]$
  - $E[g(y)|X] = \int_{-\infty}^{\infty} g(y)f(y|x)dy$
  - $E[E[X|Y, Z]|Y] = E[X|Y]$
- $Cov(X, Y) = E(XY) - E(X)E(Y)$. For $X, Y$ independent, $Cov(X, Y) = 0$
- $Cor(X, Y) : \rho_{XY} = \frac{Cov(X,Y)}{\sigma_x\sigma_y}$

- Inequalities
  - Cauchy Schwarz: $|E[XY)| \leq \sqrt{E(X^2)E(Y^2)}$
  - Markov: $P(X \geq a) \leq \frac{E(X)}{a}$ for $a > 0$
  - Chebyshev: $P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$ for $E(X) = \mu$ and $Var(X) = \sigma^2$
  - Jensen: $E[g(X)] \geq g(E[X])$ for $g$ convex; reverse if $g$ concave
    * Convex: 


- Sample Mean and Variance (iid)
  - Sample Variance: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$
  - Assume $E[x_i] = \mu$ and $Var[x_i = \sigma^2]$
    * $E(\bar{x}) = \mu$, $Var(\bar{x}) = \frac{\sigma^2}{n}$
    * $E(S^2) = \sigma^2$
- Sample Mean and Variance of Normal population (iid)
  - $\bar{x}$ and $S^2$ are independent
  - $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$
  - $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$
- Power function $\beta(\theta) = P_\theta(\boldsymbol{X} \in R)$, define rejection region
  - A test with power function $\beta(\theta)$ is **size** $\alpha$ test if $\sup_{\boldsymbol{\theta \in \Theta_0}} \beta(\theta) = \alpha$ and is **level** $\alpha$ test if $\sup_{\boldsymbol{\theta \in \Theta_0}} \beta(\theta) \leq \alpha$
- P-Value: If the rejection region is of the form $R = \{\boldsymbol{x} : T(\boldsymbol{X}) > c\}$ then the P-Value is

$$p(x) = \sup_{\boldsymbol{\theta \in \Theta_0}} P_\theta(T(\boldsymbol{X}) \geq t(\boldsymbol{x}))$$

  - NOTE: $T(\boldsymbol{X})$ is a random variable and this is the upper limit under the null hypothesis, so the statistic $t(\boldsymbol{x})$ is evaluated at the sup of the null parameter.
  - Example for standard normal with critical value $t(\boldsymbol{X}) = k_{1-\alpha}$, i.e. the upper $\alpha$ quantile critical point of standard normal, then $p(x) = P(T(\boldsymbol{X}) \geq k_{1-\alpha}) = 1 - \Phi(k_{1-\alpha})$
  - $p(x) = \sup_{\theta \in \Theta_0} P_\theta(T(X) \leq t(x))$ for rejection regions of the form $R = \{\boldsymbol{x} : T(\boldsymbol{X}) < c\}$
- UMVUE
  - Fisher information: $I_n(\theta) = -E_\theta\left(\frac{d^2}{d\theta^2}\log(L(\theta|\boldsymbol{x}))\right)$, Cramer-Rao lower bound is $\frac{[\tau'(\theta)]^2}{I_n(\theta)}$ which for i.i.d sample is $\frac{[\tau'(\theta)]^2}{nI(\theta)}$
  - Attainment: $a(\theta)[W(\boldsymbol{X}) - \tau(\theta)] = \frac{d}{d\theta}\log(L(\theta|\boldsymbol{x})$, where $W(\boldsymbol{X})$ is estimator for $\tau(\theta)$ $(E_\theta(W(\boldsymbol{X})) = \tau(\theta))$ look at $a(\theta)W(\boldsymbol{X}) - a(\theta)\tau(\theta)$
  - Lehmann-Scheffe: $g(s(\boldsymbol{x})) = E[t(\boldsymbol{x})|s(\boldsymbol{x})]$, where $g(s(\boldsymbol{x}))$ is UMVUE, $t(\boldsymbol{x})$ is an unbiased (usually crude) statistic, and is conditioned with complete sufficient statistic $s(\boldsymbol{x})$.

- Aymptotic Properties of MLE: assume regularity conditions. Let $\theta^*$ be the true value of $\theta$. For an MLE $\hat{\theta} = \hat{\theta}_n$ of $\theta$, we have that
  - (uniqueness) $\hat{\theta}_n$ is unique
  - (consistent) $\hat{\theta}_n \xrightarrow{p} \theta^*$
  - (asymptotically unbiased) $\lim\limits_{n\to\infty} E(\hat{\theta}_n) = \theta^*$
  - (asymptotically efficient) $Var(\hat{\theta}_n) \to CRLB_\theta$ as $n \to \infty$
  - (asymptotically normal) $\hat{\theta}_n \overset{asym}{\sim} N(\theta^*, CRLB_\theta)$
  - Confidence Interval using $\hat{\theta}_n$

$$P(-Z_{\frac{\alpha}{2}} \leq \frac{\hat{\theta}_n - \theta^*}{\sqrt{CRLB_\theta}} \leq Z_{\frac{\alpha}{2}}) \approx 1 - \alpha$$