Markov's Inequality
If X is a random variable that takes only nonnegative values, then, for any value $a>0$: $P(X\geq a)\leq \frac{E[X]}{a}$.

Proof: For $a>0$ let $I\begin{cases}1 & \text{if } X\geq a\\ 0 & \text{otherwise}\end{cases}$ and note that since $X\geq 0$: $I\leq \frac{X}{a}$. Taking the expectation of the preceding inequality yields: $E[I]\leq\frac{E[X]}{a}$ which, because $E[I]=P(X\geq a)$ proves the result.

Chebyshev's Inequality
If X is a r.v. with finite mean μ and variance σ^2, then for any value $k>0$: $P(|X-\mu|\geq k)\leq\frac{\sigma^2}{k^2}$

Proof: Since $(X-\mu)^2$ is a nonnegative random variable we can apply Markov's inequality. $P((X-\mu)^2\geq k^2)\leq\frac{E[(X-\mu)^2]}{k^2}$ Since $(X-\mu)^2\geq k^2$ iff $|X-\mu|\geq k$ we can also write the eq as $P(|X-\mu|\geq k)\leq\frac{E[(X-\mu)^2]}{k^2}=\frac{\sigma^2}{k^2}$ QED

CLT
Let $X_1,X_2,\ldots$ be a sequence of iid rv's with mean 0 and variance σ^2, cdf F and mgf M. Let $S_n=\sum_{i=1}^n X_i$ then $\lim_{n\to\infty}P\left(\frac{S_n}{\sigma\sqrt{n}}\leq x\right)=\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du$, cdf of $N(0,1)$

Proof: Let $Z=\frac{S_n}{\sigma\sqrt{n}}$ we will show that $M_{Z_n}(t)\xrightarrow[n\to\infty]{}e^{\frac{t^2}{2}}$. $M_{S_n}(t)=M_{X_1+X_2+\ldots+X_n}(t)=(M(t))^n$, $M_{Z_n}(t)=E[e^{tZ_n}]=E[e^{t\frac{S_n}{\sigma\sqrt{n}}}]=M_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)=\left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n$, $M(S)=M(0)+M'(0)S+\frac{1}{2}M''(0)S^2+\varepsilon$, $M'(0)=E[X_i]=0$
$M''(0)=E[X_i^2]-E^2[X_i]$ $\Rightarrow M(t)=E[e^{tX_i}]$, $M''(0)=Var[X_i]=\sigma^2$
$M_{Z_n}(t)=\left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n\approx\left(1+\frac{1}{2}\sigma^2\left(\frac{t}{\sigma\sqrt{n}}\right)^2\right)^n=\left(1+\frac{t^2}{2n}\right)^n\xrightarrow[n\to\infty]{}e^{\frac{t^2}{2}}$

Assuming ε small

Weak law of large numbers
Let $X_1,X_2,\ldots$ be a sequence of iid rv's each having $E[X_i]=\mu$. Then for any $\varepsilon>0$: $P\left(\left|\frac{X_1+\ldots+X_n}{n}-\mu\right|\geq\varepsilon\right)\to 0$ as $n\to\infty$

Proof: Additional assumption $Var[X_i]=\sigma^2$; $E\left[\frac{X_1+\ldots+X_n}{n}\right]=\mu$, $Var\left[\frac{X_1+\ldots+X_n}{n}\right]=\frac{\sigma^2}{n}$
$P\left(\left|\frac{X_1+\ldots+X_n}{n}-\mu\right|\geq\varepsilon\right)\leq\frac{\sigma^2}{n\varepsilon^2}\to 0$, when $n\to\infty$

MGF
$M_X(t)=E[e^{tX}]=\begin{cases}\sum e^{tx}\cdot P(x) & \text{if } X \text{ is discrete with density } P\\ \int_{-\infty}^{\infty}e^{tx}\cdot P(x)dx & \text{if } X \text{ is continuouty with density } P\end{cases}$

Geometric: $f(x)=(1-p)^x\cdot p$, $X=\{0,1,2,\ldots\}$
$M_X(t)=\sum e^{tx}(1-p)^x\cdot p=p\sum e^{tx}(1-p)^x$. The sum can be written as $S=1+e^t(1-p)+e^{2t}(1-p)^2+\ldots$
$S=\frac{1}{1-e^t(1-p)}$ and the mgf: $M_X(t)=\frac{pe^t}{1-(1-p)e^t}$ for $t<-\ln q$ $q=1-p$

Binomial: $f(x)=\binom{n}{k}p^k(1-p)^{n-k}$
$M_{B_{n,p}}(t)=E[e^{tk}]=\sum_{k=0}^n\binom{n}{k}p^k(1-p)^{n-k}\cdot e^{tk}=\sum_{k=0}^n\binom{n}{k}(pe^t)^k(1-p)^{n-k}=(pe^t+(1-p))^n$

Normal: $\frac{X-\mu}{\sigma}=Z\sim N(0,1)\Rightarrow X=\sigma\cdot Z+\mu\sim N(\mu,\sigma^2)$
$M_X(t)=E[e^{tX}]=E[e^{t(\sigma\cdot Z+\mu)}]=E[e^{tZ\sigma}\cdot e^{t\mu}]=e^{t\mu}\cdot E[e^{t\sigma Z}]=e^{t\mu}\cdot M_Z(t\sigma)=e^{t\mu}\cdot e^{\frac{(t\cdot\sigma)^2}{2}}=$
$e^\wedge\left(\frac{\sigma^2t^2}{2}+\mu t\right)$

utgå ifrån $\langle 1,1,1\ldots\rangle\Leftrightarrow\frac{1}{1-x}$

OGF
$G(x)=g_0+g_1x+g_2x^2+\ldots=\sum_{i=0}^{\infty}g_ix^i$

kvadrater: $\frac{d}{dx}\to\gg\to\frac{d}{dx}=\frac{1}{(1-x)^2}\to\frac{x}{(1-x)^2}\to\frac{1+x}{(1-x)^3}=\langle 1,4,9,\ldots\rangle$
kuber: $\gg\to\frac{d}{dx}=\frac{x+x^2}{(1-x)^3}\to\frac{x^2+4x+1}{(x-1)^4}=\langle 1,8,27,\ldots\rangle$
4: $\gg\to\frac{d}{dx}=\frac{x^3+4x^2+x}{(x-1)^4}\to-\frac{x^3+11x^2+11x+1}{(x-1)^5}=\langle 1,16,81,\ldots\rangle$

$\langle 1,1,1,\ldots\rangle\Leftrightarrow\frac{1}{1-x}$
$\langle 1,-1,1,-1,\ldots\rangle\Leftrightarrow\frac{1}{1+x}$
$\langle 1,0,1,0,\ldots\rangle\Leftrightarrow\frac{1}{1-x^2}$
$\langle 1,a,a^2,\ldots\rangle\Leftrightarrow\frac{1}{1-ax}$

Scaling
$\langle 1,0,1,0,\ldots\rangle\Leftrightarrow\frac{1}{1-x^2}$
$\langle 2,0,2,0,\ldots\rangle\Leftrightarrow\frac{2}{1-x^2}$

$\langle 1+x+x^2+\ldots+x^n\rangle\Leftrightarrow\frac{1-x^{n+1}}{1-x}$
"4 oranges" $=\langle 1,1,1,1,1,0\ldots\rangle=\frac{1-x^5}{1-x}$

Product rule
$G(x)\cdot F(x)=g_0\cdot f_n+g_1\cdot f_{n-1}+\ldots+g_{n-1}\cdot f_1+g_n\cdot f_0$

$E[X] = \sum x \cdot f(x)$, $E[H(X)] = \sum H(x) \cdot f(x)$, $E[aX+b] = a \cdot E[X] + b$
$Var[X] = E[X^2] - E^2[X]$

Binomial: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $E[X] = np$, $Var[X] = np(1-p)$
Geometriskt: $f(x) = (1-p)^x \cdot p$, $E[X] = \frac{1}{p}$, $Var[X] = \frac{1-p}{p^2}$
Poisson : $f(x) = (e^{-\lambda} \cdot \lambda^x) \cdot \frac{1}{x!}$, $E[X] = Var[X] = \lambda = np$
Normal : $f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{\wedge -\frac{(x-\mu)^2}{2\sigma^2}}$
Expon: $f(x) = \lambda e^{-\lambda x}$, $E[X] = \frac{1}{\lambda}$, $Var[X] = \frac{1}{\lambda^2}$

Confidence Intervals
A $100(1-\alpha)\%$ confidence interval for a parameter Θ is a random interval such that: $P[L_1 \leq \Theta \leq L_2] = 1-\alpha$

Z: $CI = \overline{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, T: $CI = \overline{X} \pm t_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

CI on σ^2: $L_1 = \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n}}$ $L_2 = \frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n}}$ $S^2 = \frac{n\sum x_i^2 - (\sum x_i)^2}{n(n-1)}$

Point estimator
$\hat{p} = \frac{X}{n} = \frac{\text{w/effect}}{n}$

CI on P
$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p}) \cdot \frac{1}{n}}$

CI on $P_1 - P_2$
$P_1 - P_2 \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$

Common OGF's
$\langle 1,0,1,\dots \rangle \Leftrightarrow \frac{1}{1-x^2} = \sum x^{2n}$, $\langle 1,2,3,\dots \rangle \Leftrightarrow \frac{1}{(1-x)^2} = \sum (n+1)x^n$, $\langle 1,3,6,10,\dots \rangle \Leftrightarrow \frac{1}{(1-x)^3} = \sum \binom{n+2}{2} x^n$

CI
Om konfidensgraden minskar, minskar även bredden på intervallet.
Om standardavvikelsen ökar ökar bredden på intervallet.
Om stickprovsstorleken minskar ökar bredden på intervallet.

How to
$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = Z \sim N(0,1)$
$P(L_1 \leq \mu \leq L_2) = (1-\alpha)$ $[L_1, L_2]$ - $100(1-\alpha)\%$ CI for μ
$(1-\alpha) = P(L_1' \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq L_2') = P(\frac{L_1'\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq \frac{L_2'\sigma}{\sqrt{n}}) = P(\frac{L_1'\sigma}{\sqrt{n}} - \overline{X} \leq -\mu \leq \frac{L_2'\sigma}{\sqrt{n}} - \overline{X}) = P(\overline{X} - \frac{L_2'\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} - \frac{L_1'\sigma}{\sqrt{n}})$

MGF: Poisson: $f(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$, $M_X(t) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} = e^{-\lambda} \sum \frac{(e^t \lambda)^x}{x!} = e^{(-\lambda + \lambda e^t)} = e^{\lambda(e^t - 1)}$

Exp: $M_X(t) = E[e^{Xt}] = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{(t-\lambda)x} dx = \lambda \left[\frac{1}{t-\lambda} \cdot e^{(t-\lambda)x}\right]_0^{\infty} =$
$= \frac{\lambda}{\lambda - t}$

Fibonacci: $f(0) = 0$
$f(1) = 1$
$f_n = f_{n-1} + f_{n-2}$

$F(x) = f_0 + f_1 x + f_2 x^2 + \dots$

$\langle 0,1,0,0,\dots \rangle \longleftrightarrow x$
$\langle 0, f_0, f_1, f_2, \dots \rangle \longleftrightarrow x \cdot F(x)$
$\langle 0, 0, f_0, f_1, \dots \rangle \longleftrightarrow x^2 F(x)$
$F(x) = \langle 0, f_0+1, f_0+f_1, \dots \rangle \longleftrightarrow x + xF(x) + x^2 F(x)$

$F(x) = x + xF(x) + x^2 F(x) \Leftrightarrow$
$F(x) = \frac{x}{1-x-x^2}$