So far: Used sample space Ω and event space $\mathcal{A}$ to describe outcome.
Now: Use a *number* to describe the outcome of an experiment.

DEF: A *random variable* x is a mapping $x : \Omega \to \mathcal{R}$ (Ω=sample space).
x associates a number with each outcome of an experiment.
EX: Flip coin 100 times. $x = (\#heads)^3$; y =#heads in 1^{st} 10 flips.

Note: Numbers represent outcomes, so sets of numbers represent events.
Q: For what subsets $F \subset \mathcal{R}$ can we compute $Pr[x \in F]$?

A1: *Induced* prob. space: $Pr[x \in F] = Pr[\omega \in \Omega : x(\omega) \in F] = Pr[x^{-1}(F)]$.
Probability space $(\Omega, \mathcal{A}, Pr)$ and random variable x define (*preimage*) *induced* probability space $(\mathcal{R}, \mathcal{A}_x, P_x)$ where $P_x[F] = Pr[x^{-1}(F)]$.

EX: Suppose $F \in \mathcal{A}_x \to P_x[F] = 0$ *or* 1. Prove x=constant with prob. 1.
Huh? Heuristically, each sample point ω of Ω maps to same constant c.

Proof: $1 = Pr[\Omega] = Pr[x^{-1}(\mathcal{R})] = Pr[x^{-1}(\cup_{n=-\infty}^{\infty}[n, n+1))]$
$= \sum_{n=-\infty}^{\infty} Pr[x^{-1}([n, n+1))] = \sum_{n=-\infty}^{\infty} P_x[[n, n+1)]$.
But: *Given* that $P_x[[n, n+1)] = 0$ *or* 1. So $P_x[[N, N+1)] = 1$ for some N.

Repeat: $1 = P_x[[N, N+1)] = P_x[[N, \frac{2N+1}{2})] + P_x[[\frac{2N+1}{2}, N+1)]$. One of two=1.
Continue: Construct *decreasing* sequence of half-open intervals $I_k = [\frac{N_k}{2^k}, \frac{N_k+1}{2^k})$
such that $I_0 \supset I_1 \supset I_2 \supset \cdots$ and $P_x[I_k] = Pr[x^{-1}(I_k)] = 1$ for all k.
Then: *Cont of prob*: $Pr[x = c] = Pr[x^{-1}(\{c\})] = P_x[\lim I_k] = \lim P_x[I_k] = 1$,
where c is the constant such that $c \in I_k$ for all k. QED.

But: Can only do this for $\{F \subset \mathcal{R} : x^{-1}(F) \in \mathcal{A}\}$=domain of Pr.
A2: If know $Pr[x \in (a, b]]$, can compute $Pr[x \in F]$ for any *Borel set F*.
i.e.: The induced probability space is $(\mathcal{R}, \mathcal{B}, P_x)$ where $\mathcal{B}$=Borel sets.

DEF: $F_x(X) = Pr[x \leq X]$=probability *distribution* function (PDF).
Then: $Pr[x \in (a, b]] = F_x(b) - F_x(a)$. Can compute $Pr[x \in F]$ for Borel sets.
Note: $F_x(X)$ or $P_x(X)$ or $F_X(x)$ or "cumulative dist. function" (CDF).

Props: $F_x(X)$ is *nondecreasing*: $x_1 < x_2 \to F_x(x_1) \leq F_x(x_2)$ (may be level).
$\lim_{X \to -\infty} F_x(X) = Pr[x \leq -\infty] = 0$; $\lim_{X \to \infty} F_x(X) = Pr[x \leq \infty] = 1$.
$F_x(X)$ is *continuous from the right*: $\lim_{\epsilon \to 0^+} F_x(X + \epsilon) = F_x(X)$.

Q1: In how many different ways can N objects be ordered?
A1: $N! = N(N-1)(N-2)\cdots(2)(1) \simeq N^N e^{-N}\sqrt{2\pi N}$ (Stirling's formula).

Q2: In how many ways can K objects chosen from N be ordered?
A2: $P_K^N = N(N-1)(N-2)\cdots(N-K+1) = N!/(N-K)!$ (choose K).

Q3: In how many ways can K objects be chosen from N *without* ordering?
A3: No longer want K! ordering of the K objects chosen, so $C_K^N = \frac{1}{K!}P_K^N$.
$C_K^N = N(N-1)\cdots(N-K+1)/K! = \frac{N!}{K!(N-K)!} = \binom{N}{K} = \binom{N}{N-K}$.

Q4: A class of 88 students is divided into 4 recitations of 22 students each. In how many ways can this be done?
A4: $\binom{88}{22}\binom{66}{22}\binom{44}{22}\binom{22}{22} = \frac{88!}{(22!)^4} = 1.162 \times 10^{50}$.

Q5: Multinomial formula: $(\sum_{i=1}^{M} a_i)^N = \underbrace{\sum_{i_1}\cdots\sum_{i_{M-1}}}_{i_1+\cdots+i_M=N} \frac{N!}{i_1!\cdots i_M!} a_1^{i_1}\ldots a_M^{i_M}$.
A5: See Stark and Woods pp. 22-36 for a nice treatment of combinatorics.

Given: An urn contains 3 red balls and 7 black balls (10 balls total). What is the probability that 2 of 6 balls picked out of the urn are red?

Q1: If balls are picked *without* replacement (simultaneously), use the *hypergeometric* formula: $Pr = \binom{3}{2}\binom{7}{4}/\binom{10}{6}$.

Q2: If balls are picked *with* replacement (one at a time), use the *binomial* formula: $Pr = \binom{6}{2}\left(\frac{3}{10}\right)^2\left(\frac{7}{10}\right)^4$.

Δ: Suppose first ball picked is red. Pr[second ball picked is red]=? **with** replacement: still $\frac{3}{10}$. **without** replacement: now $\frac{2}{9}$.
Then: This leads to hypergeometric formula. Note multiplication commutes.

DEF: $f_x(X) = \frac{dF_x(X)}{dX}$=probability *density* function (pdf) (vs. distribution).
Then: $Pr[a < x \leq b] = \int_a^b f_x(X)dX$. pdf exists only if PDF differentiable.
Note: As $\delta X \to 0$, $Pr[X < x \leq X + \delta X] = f_x(X)\delta X$. $f_x(X)$ NOT a prob!

Props: $f_x(X) \geq 0; > 1$ OK; $\int_{-\infty}^{\infty} f_x(X)dX = 1$; $Pr[A] = \int_A f_x(X)dX$.
Note: $Pr[x = b] \neq 0 \to f_x(X)$ has impulse at $x = b$; include in $\int_a^b f_x(X)dX$.