

	Distro	E(X)	V(X)	pmf
Binomial (discrete)	$X \sim \text{Bin}(n, p)$	np	$np(1-p)$	$\binom{n}{x} p^x (1-p)^{n-x}$
Hypergeo (discrete)	$X \sim \text{Hyper}(n, M, N)$	$n \cdot \frac{M}{N}$	$\left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$	$\frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$
NegBin (discrete)	$X \sim \text{NegBin}(r, p)$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\binom{x+r-1}{r-1} p^r (1-p)^x$
Poisson (discrete)	$X \sim \text{Poi}(\mu)$	μ	μ	$\frac{e^{-\mu} \cdot \mu^x}{x!}$

$\binom{n}{k} = \frac{n!}{r!(n-r)!}$ = “n choose k” = num combos of size r that can be formed from n indiv’s in group **conditions for Bin (1)** Two possible outcomes of each trial: S (what we counting) or F (all else). (2) num trials n is

known and fixed (3) outcomes are independent from one trial to others (4) p (success) is same for all trials (·) if conditions met, X = num S’s is binomial RV w/ params n number of trials / sample size and p probability of success. **conditions for HyperG** when sampling *without replacement* from a small population (population not at least $10 \times$ sample size). (1) pop is finite, consists of N individuals/objects (2) 2 kinds of indivs/objs: success (S), failure (F); and there are $M < N$ successes in population (3) a sample of n individuals are randomly selected without replacement (·) if conditions met, X = num S’s is hypergeom rv **pmf explained** numerator: num ways to have x successes & $n-x$ failures. Denominator: num ways to select a subset of n indivs/objs out of group of N **Finite Pop** **Correction Factor** = $1 - [\frac{n-1}{N-1}]$, second term is proportion of population included in sample. Since < 1 , hyperG has *smaller* $V(X)$ than bin. **conditions for NegBin (1)** exp consists of *potentially infinite num of indep Bernoulli trials* (2) two poss outcomes for each trial: S, F (3) probability of success p is *same for all trials* (4) exp continues until a total of r successes have been observed, where r is fixed int > 0 (·) if conditions met, X = num fails before r th success is negbin rv dist dependent on r, p **Geom dist** special case of NegBin where $r = 1$ **$P(12 \leq X \leq 28) = P(11 < X < 29) = P(X = 12) + \dots + P(X = 28)$** **Continuous RVs** change: *integration* instead of summation is now used to calculate probs, E’s, V’s. Also, CDF’s play more central role in finding probs, *percentiles* **conditions for continuity** X is continuous if (i) X can be any value in an intrvl, such as $[0,1]$ or even $(-\infty, \infty)$, or union of disjoint intrvls, AND (ii) $P(X = c) = 0$ for any possible val of c —i.e., no indiv. val of X has a positive probability (·) CRV’s often used to represent measurements **Probability Density Function (pdf)** $f(x)$ defines shape of a continuous distro. Needs to be integrated to find probabilities. Has prop that, for a crv X has property that, for any two constants a, b : $P(a \leq X \leq b) = \int_a^b f(x)dx$ **pdf criteria (1)** $f(x)$ is *non-negative*: $f(x) \geq 0$ for all $-\infty < x < \infty$ (2) entire area under $f(x)$ is 1: $\int_{-\infty}^{\infty} f(x)dx = 1$ **$P(X \leq 0.5) = \int_a^{0.5} f(x)dx$** **Uniform Dist** $f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$, 0 otherwise (issa square wave) **Cumulative Distribution Function (cdf)**: the cdf $F(x)$ of a crv X is $F(x) = P(X \leq x) = \int_{-\infty}^x f(y)dy$ **cdf of uniform dist** $F(x) = \int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a}$ ← straight line w/ $F(a) = 0, F(b) = 1$ and slope $1/(b-a)$ **CDF Propositions** if X is a crv with pdf $f(x)$ and cdf $F(x)$, a, b any two numbers s.t. $a < b$, then (1) $P(X > a) = 1 - F(a)$ (2) $P(a \leq X \leq b) = F(b) - F(a)$ (3) $P(X = x) = 0$ for all values of x . Resultantly, (3.i) $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$ (4) $f(x)$ is the derivative of $F(X)$: $f(x) = \frac{dF(x)}{dx}$ **Expected value of X , $\mu_X = E(X) = \int x \cdot f(x)dx$** **Expected value of function, $\mu_{h(X)} = E(h(x)) = \int h(x) \cdot f(x)dx$** **Variance of X , $\sigma_X^2 = V(X) = \int (x - \mu)^2 \cdot f(x)dx = E(X^2) - [E(X)]^2$** **Standard Deviation of X , $\sigma_X = \sqrt{V(X)}$** **Percentile** for any continuous rv X , and for values of p ($0 < p < 1$), the $(100p)^{th}$ percentile, $\eta(p)$, of dist X is defined by $p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)dy$ **notes on percentile** $\eta(p)$ is the val in the distribution of X s.t. $100p\%$ of area under density curve lies to the left of $\eta(p)$ and $100(1-p)\%$ lies to the right of $\eta(p)$. **to find p^{th} percentile of any dist**: set p equal to the cdf and solve for x ! **Median $\tilde{\mu}$ of X** for any crv X , the median is the 50^{th} percentile of the dist X , s.t. $\tilde{\mu}$ satisfies $0.5 = F(\tilde{\mu})$. Half of area under density curve lies to left of $\tilde{\mu}$, and half to right. **mean vs median** median is middle val, won’t be affected by outlier values. However, mean *will* be affected. If distribution is right-skewed, median $<$ mean; if left-skewed, mean $<$ median **Normal (Gaussian) Dist, $X \sim N(\mu, \sigma^2)$** most common family; many rv’s have (approx) normal dist, and if they don’t, ther mean often does. Most large-sampling statistic inference relies upon normal distribution. **Normal pdf** a crv X has a normal distribution w/ parameters μ and σ (or μ and σ^2) if pdf of X is $f(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$ **notes on normal dist (1)** x can take any vals $-\infty < x < \infty$ (2) μ (although fixed) can take any vals $-\infty < x < \infty$ (2.1) μ is the *mean* of the normal dist; this controls center of distribution (3) σ can take any vals $\sigma > 0$ (3.1) σ is the *standard deviation* of the normal dist (σ^2 is the variance); this controls the spread/scale of the dist (·) normal dist always has bell shape **sketching a normal dist (i)** curve is tallest at μ , at center (ii) 99.7% of dist is between $\mu - 3\sigma$ and $\mu + 3\sigma$. These are the approx. smallest and largest vals of X (iii) changing μ shifts entire distribution (iv) changing σ controls how tall/flat dist is **standard score $z = \frac{x-\mu}{\sigma}$** tells you how many *standard deviations* a particular observation is from mean. Dist from mean relative to how much *variation* there is. Follows a *standard normal distribution* **standard normal distribution** a crv Z is said to have a *standard normal distribution* if $\mu = 0$ and $\sigma = 1$: $Z \sim N(0, 1)$ **pdf $Z = f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$** **cdf $Z \Phi(z) = P(Z \leq z) = \int_{-\infty}^z f(y; 0, 1)dy$** **can find probabilities by (1)** integrating the pdf $f(x)$, so $P(a < X < b) = \int_a^b f(x)dx$ (2) plugging values into cdf $F(x)$, so $P(a < x < b) = F(b) - F(a)$ **finding probabilities using a normal dist (i)** integrals must be approximated; difficult before computers were available (ii) lots of possible normal distributions (iii) soluiton (then): “standardize” each normal dist using standard score; get probabilities from table (iv) solution (now): computing power **two ways to calc probabilities for a normal dist: (1)** use table A3 in appendix, which gives values for $\Phi(z)$ (2) use graphing calc **e.g.** for $Z \sim N(0, 1)$, $P(Z < -1.37) = \Phi(-1.37)$, so nav to row $z = -1.3$ and col $z = 0.07$ and it’s there **for greater-thans**: table only gives CDF = $P(Z < z)$. To find $P(Z > z)$, need to subtract prob provided by table from 1 **inequalities** $P(-2 < Z < 2) = \Phi(2) - \Phi(-2)$ (see **pic 1, pic 3**) **calc** TI-83: $2^{nd} \rightarrow \text{VARs} \rightarrow 2:\text{normalcdf}(\text{LB}, \text{UB}, \mu, \sigma)$. if one of bounds is infinity, enter really big (/small) number. to find **percentile** via calc, use **invNorm** func. **un-standardizing a table val** $x = z \cdot \sigma + \mu$ **standardizing a probability** see **pic 2, pic 4** **Normal approx. to Bin** $X \sim \text{Bin}(n, p)$ can be written as $X = X_1 + \dots + X_n$ **Central Limit Thm**: This X is approximately normal or $Z = \frac{X - np}{\sqrt{npq}}$ is approx standard normal $N(0, 1)$ as $n \rightarrow \infty$. Good approx even if n moderate (≈ 30) and $np \geq 5$ and $n(1-p) \geq 5$ **Standardizing Normal Dists** if not standard normal dist, you’ll have to standardize X val before using table **$B(n, p) \approx N(np, npq)$** **ex** $X \sim \text{Bin}(n = 15, p = 0.4)$, $Y \sim N(np, npq) = N(6, 3.6)$ then $P(X = 5) \approx P(4.5 < Y < 5.5) = P(\frac{4.5-6}{\sqrt{3.6}} < P(\frac{Y-6}{\sqrt{3.6}} < P(\frac{5.5-6}{\sqrt{3.6}}) = P(Z < -0.2635) - P(Z < -0.7906) = 0.1815$, while exact $P(X = 5) = 0.1859$ **Gamma function $\Gamma(\alpha)$** (1) for any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$. (2) For any positive int n , $\Gamma(n) = (n - 1)!$ (3)

$\Gamma(1/2) = \sqrt{\pi}$ **Gamma Distribution** $X \sim \text{Gamma}(\alpha, \beta)$ not the gamma function!!! Often used for waiting times or survival times
Gamma PDF $f(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}$ **Gamma CDF** $F\left(\frac{x}{\beta}; \alpha\right) = \int_0^{x/\beta} \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy$ (probabilities found via a table) **Gamma E(X)**
 $= \alpha\beta$ **Gamma V(X)** $= \sigma^2 = \alpha\beta^2$ **Gamma props** (1) x can only take non-neg vals: $x \geq 0$ (2) β can only take pos vals: $\beta > 0$.
 β is the scale parameter; it controls spread of dist. When $\beta = 1$, X is a standard gamma dist: $f(x; \alpha) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}$ **Gamma prob**
 $P(X < 6) = F(6/1; 3)$ go to table A4, look for col $\alpha = 3$ and down to row $X = 6$. $P(2 < X < 5) = F(5/1; 3) - F(2/1; 3)$ **Standardizing Gamma Dists** if $\beta \neq 1$, standardize X val before looking it up in table. e.g., for $G(4, 5)$, $P(X < 8) = F(8/5; 4) = F(1.6; 4)$, make inequality: $F(1; 4) < F(1.6; 4) < F(2; 4) \rightarrow 0.019 < P(X < 8) < 0.143$ **Special cases of Gamma dist** (i) when $\alpha = 1$ and $\beta > 0$, gamma dist is the Exponential dist w/ parameter β (ii) when $\alpha = v/2$ and $\beta = 2$, for pos int v , gamma dist is Chi-square dist w/ param v —common dist for stat inference **Exponential Dist** $X \sim \text{Exp}(\lambda)$ special case of gamma dist. x is non-neg, λ is positive. Gen. skewed to right **Exponential PDF** $f(x; \lambda) = \lambda e^{-\lambda x}$ **Exponential CDF** $F(x) = 1 - e^{-\lambda x}$ **Exponential E(X)** $= \frac{1}{\lambda}$ **Exponential V(X)** $= \frac{1}{\lambda^2}$ **hazard rate** λ higher value of λ means shorter survival times. “Waiting time” could also represent length of time between successes of a Poisson rv **notes Exp** (i) sum of α pos independent $\text{Exp}(\beta)$ rv’s can be modeled by the $\text{Gamma}(\alpha, \beta)$ dist (ii) exp dist is said to be memoryless, i.e. $P(X > t + s | X > t) = P(X > s)$ (iii) gamma and other dists don’t have this property **system of components** each component lifetime $\sim \text{Exp}(0.01)$, component failures are independent, A_i = i th component lasts at least t hours, and X = time at which sys fails. Compute cdf X : $F(t) = P(X \leq t)$. Start by finding $P(X > t)$: $X > t \Leftrightarrow A_1 \cap A_2 \cap A_3 \cap A_4$, so $P(X > t) = P(A_1) \cdots P(A_4) = P(Y_1 > t) \cdots P(Y_4 > t)$ where Y_i = lifetime of component i . Note cdf Y_i is $P(Y_i \leq t) = (1 - e^{-0.01t})$, so $P(Y_i > t)$ is 1 minus that $= e^{-0.01t}$. Then $P(X > t) = e^{-0.04t} \Rightarrow F(t) = P(X \leq t) = 1 - e^{-0.04t}$, which is $X \sim \text{Exp}(0.04)$ **other exs** lightbulb follows exp dist; expected lifetime 1,000 hrs. Prob rand-select bulb will last $> 2,000$ hrs is, since we can get $\lambda = 0.001$ from $E(X)$, $P(X > 2000) = 1 - F(2000; 0.001) = e^{-(0.001)(2000)} = 0.135$ **practice probs NegBin ex** An expert sharpshooter misses a target 5% of the time. Let X = the number of shots that hit the target before the second miss. What is the probability that she misses the target for the second time on the fifteenth shot? How many shots would you expect to hit the target before the second miss? How many shots would you expect her to take before the second miss? **NegBin ex** $X \sim \text{NegBin}(r = 2, p = 0.05)$, so (a) required prob $= P(X = 15 - r = 13) = \binom{14}{1} p^2 (1 - p)^{13} = 14 \cdot 0.05^2 \cdot 0.95^{13} = 0.018$ (b) $E(X) = \frac{r(1-p)}{p} = \frac{2(1-0.05)}{0.05} = 38$ (c) $E(X) + 2 = 40$ **HyperGeo ex** Among 16 applicants for a job, 11 have college degrees. Suppose that 4 of these applicants will be randomly selected for an interview. Let X = the number of selected applicants who have a college degree. What is the distribution of X ? What is the probability that at least 1 of the applicants has a college degree? **HyperGeo ex** (a) $X \sim \text{hypergeo}(N = 16, M = 11, n = 4)$ (b) $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{11}{0} \binom{5}{4}}{\binom{16}{4}} = 1 - \frac{5}{1820} = 0.9973$ (c) **Poisson ex** Suppose that earthquakes occur in the western portion of the United States at an average rate of 2 per week. Let X = the number of earthquakes that occur in the western portion of the United States in a week. What is the probability that at most 1 earthquake occurs during the next week? **Poisson ex** (\cdot) $X \sim \text{Poisson}(\mu = 2)$, so $P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{\mu^0 e^{-\mu}}{0!} + \frac{\mu^1 e^{-\mu}}{1!} = (1 + 2)e^{-2} = 0.406$ **Normal ex** In a photographic process, the developing time of prints may be looked upon as a normal random variable with a mean of 18 seconds and a standard deviation of 1.2 seconds. What is the probability that it takes between 15.9 and 19.3 seconds to develop a print? Fill in the blank: 85% of all prints are developed in _____ seconds or less. **Normal ex** (\cdot) X = developing time $\sim \text{Normal}(\mu = 18 \text{ sec}, \sigma = 1.2 \text{ sec})$, so (a) $P(15.9 \leq X \leq 19.3) = P(-1.75 \leq Z \leq 1.08) = 0.82$, where we standardized by plugging in $Z = \frac{(X - \mu)}{\sigma}$ for both bounding X -vals and, in the middle, in the general case. (b) 85 percentile from $N(0, 1)$ is 1.04, so required number is $18 + 1.04 \cdot 1.2 = 19.25$ seconds **Uniform ex** The time taken to deliver a pizza follows a uniform distribution from 20 minutes to 60 minutes. What is the probability that the time to deliver a pizza is at least 25 minutes? What is the 55th percentile of this distribution? **Uniform ex** (a) $X \sim \text{Uniform}(20, 60)$, so $P(X \geq 25) = (60 - 25)/(60 - 20) = 0.875$ (b) $(x - 20)/(60 - 20) = 0.55 \Rightarrow x = 42$ minutes **Exponential ex** The lifetime of a certain type of car (in thousands of miles) has an Exponential Distribution with $\lambda = 0.01$. What is the probability that a car of this type last for more than 90,000 miles? **Exponential ex** (a) $X \sim \text{exp}(\lambda = 0.01)$, so $P(X > 90) = e^{-90\lambda} = e^{-0.9} = 0.4066$ **Binomial ex** The College Board reports that 2% of the 2 million high school students who take the SAT each year receive special accommodations because of documented disabilities (Los Angeles Times, July 16, 2002). Consider a random sample of 25 students who have recently taken the test. **Binomial ex** (a) $P(X = 1) = \binom{25}{1} p^1 (1 - p)^{25-1} = 25 \cdot 0.02 \cdot (1 - 0.02)^{24} = 0.3079$ (b) $P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - 0.02)^{25} = 0.3965$ (c) $P(X \geq 2) = 1 - P(X \leq 1) = 1 - (P(X = 0) + P(X = 1)) = 1 - (0.6034 + 0.3079) = 0.3079$ (d) $\mu = np = 25 \times 0.02 = 0.5$, $\sigma = \sqrt{np(1 - p)} = \sqrt{25 \times 0.02 \times (1 - 0.02)} = 0.7$, so $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-0.9 \leq X \leq 1.9) = P(X \leq 1) = 0.9113$ (e) Y = hours a randomly selected student is allowed. Then Y takes 4.5 with prob=0.02 and 3 with prob=0.98. $E(Y) = 4.5 \cdot 0.02 + 3 \cdot 0.98 = 3.03$ hours **Binomial Func ex** A toll bridge charges \$1.00 for passenger cars and \$2.50 for other vehicles. Suppose that during daytime hours, 60% of all vehicles are passenger cars. If 25 vehicles cross the bridge during a particular daytime period, what is the resulting expected toll revenue? (Hint: Let X = the number of passenger cars; then the toll revenue $h(X)$ is a linear function of X .) **Binomial Fun ex** (\cdot) X = # of passenger cars out of $n = 25$. Then $X \sim \text{Bin}(n = 25, p = 0.6)$. The toll $Y = X + 2.5(n - X)$, so $E(Y) = E(X) + 2.5(n - E(X)) = np + 2.5(n - np) = 15 + 2.5(25 - 15) = \40 **cdf ex** The weekly demand for propane gas (in 1000s of gallons) from a particular facility is an rv X with pdf $f(x) = 2(1 - \frac{1}{x^2})$, $1 \leq x \leq 2$, 0 otherwise. (a) Compute the cdf of X , (b) Obtain an expression for (100p)th percentile. What is the value of $\tilde{\mu}$? (c) Compute $E(X)$ and $V(X)$. (d) If 1.5 thousand gallons are in stock at the beginning of the week and no new supply is due in during the week, how much of the 1.5 thousand gallons is expected to be left at the end of the week? (Hint: let $h(x)$ =amount left when demand = x) **cdf ex** (a) $F(X) = 0$ for $x < 1$ and $F(x) = 1$ for $x > 2$. For x in $[1, 2]$, $F(x)$ is equal to $F(x) = \int_1^x 2(1 - \frac{1}{y^2}) dy = 2(x - 2 + x^{-1})$ (b) Solve $2(x - 2 + x^{-1}) = p$ for x in $(1, 2) \Rightarrow x = (4 + p + \sqrt{p^2 + 8p})/4$, so set $p = 0.5 \Rightarrow \mu = (4.5 + \sqrt{4.5})/4 = 1.64$ (c) $E(X) = \int_1^2 2x(1 - x^{-2}) dx = \int_1^2 2x dx - 2 \int_1^2 x^{-1} dx = x^2|_1^2 - 2 \ln(x)|_1^2 = 3 - 2 \ln(2)$ (d) Let $Y = h(X)$ be the left amt. Then $h(x) = 1.5 - x$ if $x < 1.5$, otherwise $h(x) = 0$. So $E(Y) = E(h(X)) = \int_1^{1.5} h(x) f(x) dx = \int_1^{1.5} (1.5 - x) 2(1 - x^{-2}) dx = 2 \int_1^{1.5} (1.5 - x - 1.5/x^2 + 1/x) dx = 2(1.5 \times 0.5 - (1.5^2 - 1)/2 - 1.5 \times (1 - 1/1.5) + \ln 1.5) = 0.061$

$$\mathbf{V(X)}= E(X^2) - (E(X))^2 \text{ } \sigma =\sqrt{V(X)}$$