	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* $(\cdot)$ 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× 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* $(\cdot)$ if conditions met, X = num S's is hypergeom rv **pmf explained** numerator: num ways to have x successes & $n-x$ failues. 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 $(\cdot)$ if conditions met, X = num fails before rth success is negbin rv dist dependent on r,p **Geom dist** special case of NegBin where $r=1$ $P(12\leq X\leq 28)=P(11<X<29)=P(X=12)+\ldots+P(X=28)$ **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 (-∞,∞), 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 $(\cdot)$ CRV's often used to represent measurements **Probability Density Funciton (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: $\int_{-\infty}^{\infty}f(x)dx=1$ $\mathbf{P(X\leq 0.5)}=\int_a^{0.5}f(x)dx$ **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** $F(x)=\int_a^x\frac{1}{b-a}dy=\frac{x-a}{b-a}\leftarrow$ 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**, $\mu_X=E(X)=\int_x x\cdot f(x)dx$ **Expected value of function,** $\mu_{h(X)}=E(h(x))=\int_x h(x)\cdot f(x)dx$ **Variance of X,** $\sigma_X^2=V(X)=\int_x(x-\mu)^2\cdot f(x)dx=E(X^2)-[E(X)]^2$ **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,** $X\sim\mathbf{N}(\mu,\sigma^2)$ 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 $(\cdot)$ 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 $\sigma=1$: $Z\sim\text{N}(0,1)$ **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 $P(a<X<b)=\int_b^a f(x)dx$ (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 $Z\sim\text{N}(0,1)$, $\text{P}(Z<-1.37)=\Phi(-1.37)$, 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} →VARS→2:normalcdf(LB,UB,μ,σ). 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 $X=X_1+\ldots+X_n$ **Central Limit Thm:** This X is approximately normal or $Z=\frac{X-np}{\sqrt{npq}}$ is approx standard normal $N(0,1)$ as $n\to\infty$. 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** $X\sim\text{Bin}(n=15,p=0.4)$, $Y\sim\text{N}(np,npq)$ =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** $X\sim\mathbf{Gamma}(\alpha,\beta)$ *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: $F(1;4)<F(1.6;4)<F(2;4)\to 0.019<P(X<8)<0.143$ **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** $X\sim\mathbf{Exp}(\lambda)$ 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 sucsesses of a *Poisson* rv **notes Exp** (i) sum of α pos independent Exp(β) rv's can be modeled by the Gamma(α, β) 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 comopnents** each component lifetime $\sim$Exp(0.01), component failures are independent, A_i =ith component lasts at least t hours, and X = time at which sys fails. Compute cdf X: $F(t) = P(X \le t)$. 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 $P(Y_i \le t) = (1 - e^{-0.01t})$, so $P(Y_i > t)$ is 1 minus that $= e^{-0.01t}$. Then $P(X > t) = e^{-0.04t} \Rightarrow F(t) = P(X \le t) = 1 - e^{-0.04t}$, which is $X \sim$Exp(0.04) **other exs** lifebulb 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 $E(X)$, $P(X > 2000) = 1 - F(2000; 0.001) = e^{-(0.001)(2000)} = 0.135$

Ex: $X \sim$ N(7,4)

- Use Table A.3 to find $P(6 < X < 11)$.

$$P(6 < X < 11) = P\left(\frac{6-7}{2} < \frac{X-7}{2} < \frac{11-7}{2}\right)$$
$$= P(-0.5 < Z < 2)$$
$$= \Phi(2) - \Phi(-0.5)$$
$$= 0.9772 - 0.3085$$
$$= 0.6687$$

Ex: Scores ~ N(75, 25). Top 20% = excellent; bottom 25% = in danger; bottom 5% = failing

a. What is the probability that a randomly selected school will score below 70?

$$P(X < 70) = P\left(\frac{X-75}{5} < \frac{70-75}{5}\right)$$
$$= P(Z < -1) = 0.1587$$

There is a 15.87% chance of a school scoring below 70 on this test.

Ex: Scores ~ N(75, 25). Top 20% = excellent; bottom 25% = in danger; bottom 5% = failing

c. What is the probability that a randomly selected school will score between 85 and 90?

$$P(85 < X < 90) = P\left(\frac{85-75}{5} < \frac{X-75}{5} < \frac{90-75}{5}\right)$$
$$= P(Z < 3) - P(Z < 2)$$
$$= 0.9987 - 0.9772 = 0.0215$$

There is a 2.15% chance of a school scoring between 85 and 90 points on this test.

Ex: Scores ~ N(75, 25). Top 20% = excellent; bottom 25% = in danger; bottom 5% = failing

d. What is the score cut-off required for schools to be labeled excellent? Show all work.

Want upper 20%, which corresponds to 80th percentile

Look in middle of Table A.3: The z-score with a probability closest to 0.8000 is z = 0.84; we need to un-standardize this z value:

z = (x – μ) / σ

0.84 = (x – 75) / 5

0.84(5) = x – 75

x = 0.84(5) + 75 = 79.2

The cut-off score for schools to be labeled "excellent" is 79.2 points; 80% of all schools score below this value.

78. According to the article **"Characterizing the Severity and Risk of Drought in the Poudre River, Colorado"** (***J. of Water Res. Planning and Mgmnt.*, 2005: 383–393**), the drought length Y is the number of consecutive time intervals in which the water supply remains below a critical value y_0 (a deficit), preceded by and followed by periods in which the supply exceeds this critical value (a surplus). The cited paper proposes a geometric distribution with $p = .409$ for this random variable.

a. What is the probability that a drought lasts exactly 3 intervals? At most 3 intervals?

b. What is the probability that the length of a drought exceeds its mean value by at least one standard deviation?

Y has a negative binomial distribution with $r = 1$ and $p = 0.409$.

(a) $P(Y = 3) = 0.409(1 - 0.409)^3 = 0.0844$,

$P(Y \le 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) = 0.409(1 - 0.409)^0 + 0.409(1 - 0.409)^1 + 0.409(1 - 0.409)^2 + 0.409(1 - 0.409)^3 = 0.8780.$

(b) $E(Y) = (1 - 0.409)/0.409 = 1.4450$,

$V(Y) = (1 - 0.409)/0.409^2 = 3.5330$,

and $\sigma = \sqrt{V(Y)} = 1.8796$.

Thus $P(Y > 1.4450 + 1.8796) = P(Y > 2) = 1 - P(Y = 0) - P(Y = 1) = 0.3493$.

81. Suppose that the number of drivers who travel between a particular origin and destination during a designated time period has a Poisson distribution with parameter $\mu = 20$ (suggested in the article **"Dynamic Ride Sharing: Theory and Practice,"** ***J. of Transp. Engr.*, 1997: 308–312**). What is the probability that the number of drivers will

a. Be at most 10?

b. Exceed 20?

c. Be between 10 and 20, inclusive? Be strictly between 10 and 20?

d. Be within 2 standard deviations of the mean value?

(a) $P(X \le 10) = F(10, 20) = 0.11$.

(b) $P(X > 20) = 1 - P(X \le 20) = 1 - F(20, 20) = 1 - 0.559 = 0.441$.

(c) $P(10 \le X \le 20) = P(9 < X \le 20) = F(20, 20) - F(9, 20) = 0.559 - 0.005 = 0.554$.

$P(10 < X < 20) = P(10 < X \le 19) = F(19, 20) - F(10, 20) = 0.470 - 0.011 = 0.459$.

(d) $P(\mu - 2\sigma \le X \le \mu + 2\sigma) = P(20 - 2\sqrt{20} \le X \le 20 + 2\sqrt{20}) = P(11.06 \le X \le 28.94) = P(11 < X \le 28) = F(28, 20) - F(11, 20) = 0.966 - 0.021 = 0.945$.

6. The actual tracking weight of a stereo cartridge that is set to track at 3 g on a particular changer can be regarded as a continuous rv X with pdf

$$f(x) = \begin{cases} k[1 - (x-3)^2] & 2 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

a. Sketch the graph of $f(x)$.

b. Find the value of k.

c. What is the probability that the actual tracking weight is greater than the prescribed weight?

d. What is the probability that the actual weight is within .25 g of the prescribed weight?

e. What is the probability that the actual weight differs from the prescribed weight by more than .5 g?

(b) $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_2^4 k[1 - (x-3)^2]\mathrm{d}x = 4k/3 = 1$. Therefore, $k = 0.75$.

(c) $P(X > 3) = \int_3^4 0.75[1 - (x-3)^2]\mathrm{d}x = 0.5$.

(d) $P(2.75 \le X \le 3.25) = \int_{2.75}^{3.25} 0.75[1 - (x-3)^]\mathrm{d}x = 0.367$.

(e) $P(X < 2.5) + P(X > 3.5) = 1 - P(2.5 \le X \le 3.5) = 1 - \int_{2.5}^{3.5} 0.75[1 - (x-3)^]\mathrm{d}x = 0.3125$.

22. The weekly demand for propane gas (in 1000s of gallons) from a particular facility is an rv X with pdf

$$f(x) = \begin{cases} 2\left(1 - \dfrac{1}{x^2}\right) & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

a. Compute the cdf of X.

b. Obtain an expression for the $(100p)$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.]

(a) For $1 \le x \le 2$, $F(x) = \int_1^x 2(1 - \frac{1}{x^2})\mathrm{d}x = 2x + \frac{2}{x} - 4$. Thus,

$$F(x) = \begin{cases} 0, & \text{if } x < 1 \\ 2x + \frac{2}{x} - 4, & \text{if } 1 \le x \le 2 \\ 1, & \text{if } x > 2 \end{cases}$$

(b) $2x + \frac{2}{x} - 4 = p$, which gives $2x^2 - (4+p)x + 2 = 0$, or equivalently, $(x - p/4 - 1)^2 = (p/4+1)^2 - 1 = \frac{1}{16}(p^2 + 8p)$. Solving the equation, we get $x = \frac{p+4+\sqrt{p^2+8p}}{4}$. When $p = 1/2$, we get $\tilde{\mu} = \frac{.5+4+\sqrt{.5^2+8(.5)}}{4} = 1.640$.

(c) $E(X) = \int_1^2 2x(1 - \frac{1}{x^2})\mathrm{d}x = (x^2 - 2\log(x))|_1^2 = (2^2 - 2\log(2)) - (1^2 - 2\log(1)) = 3 - 2\log(2) = 1.614$.

$E(X^2) = \int_1^2 2x^2(1 - \frac{1}{x^2})\mathrm{d}x = (\frac{2x^3}{3} - 2x)|_1^2 = (16/3 - 2(2)) - (2/3 - 2(1)) = 2.667$.

$V(X) = E(X^2) - (E(X))^2 = 2.667 - 1.614^2 = 0.063$.

(d) $E(X) = 1.614 > 1.5$, hence none.

11. Let X denote the amount of time a book on two-hour reserve is actually checked out, and suppose the cdf is

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2}{4} & 0 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

a. Calculate $P(X \le 1)$.

b. Calculate $P(.5 \le X \le 1)$.

c. Calculate $P(X > 1.5)$.

d. What is the median checkout duration $\tilde{\mu}$? [solve $.5 = F(\tilde{\mu})$].

e. Obtain the density function $f(x)$.

f. Calculate $E(X)$.

g. Calculate $V(X)$ and σ_X.

h. If the borrower is charged an amount $h(X) = X^2$ when checkout duration is X, compute the expected charge $E[h(X)]$.

(a) $P(X \le 1) = F(1) = \frac{1^2}{4} = .25$.

(b) $P(.5 \le X \le 1) = F(1) - F(.5) = \frac{1^2}{4} - \frac{.5^2}{4} = .1875$.

(c) $P(X > 1.5) = 1 - P(X \le 1.5) = 1 - F(1.5) = 1 - \frac{1.5^2}{4} = .4375$.

(d) $.5 = F(\tilde{\mu}) = \frac{\tilde{\mu}^2}{4}$, which gives $\tilde{\mu}^2 = 2$ and $\tilde{\mu} = \sqrt{2} = 1.414$.

(e) $f(x) = F'(x) = \frac{x}{2}$ for $0 \le x < 2$ and 0 otherwise.

(f) $E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^2 x\frac{x^2}{2} = \frac{x^3}{6}|_0^2 = 1.333$.

(g) $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^2 x^2\frac{x^2}{2} = \frac{x^4}{8}|_0^2 = 2$. Thus, $V(X) = E(X^2) - (E(X))^2 = 2 - 1.333^2 = .222$. Hence, $\sigma_X = \sqrt{.222} = .471$.

(h) From g, $E(X^2) = 2$.

38. There are two machines available for cutting corks intended for use in wine bottles. The first produces corks with diameters that are normally distributed with mean 3 cm and standard deviation .1 cm. The second machine produces corks with diameters that have a normal distribution with mean 3.04 cm and standard deviation .02 cm. Acceptable corks have diameters between 2.9 cm and 3.1 cm. Which machine is more likely to produce an acceptable cork?

Page 168, Problem 38

$P(2.9 < X_1 < 3.1) = P(\frac{2.9-3}{.1} < \frac{X_1-3}{.1} < \frac{3.1-3}{.1}) = P(-1 < Z < 1) = P(Z < 1) - P(Z < -1) = .8413 - .1587 = .6826$.

$P(2.9 < X_2 < 3.1) = P(\frac{2.9-3.04}{.02} < \frac{X_2-3.04}{.02} < \frac{3.1-3.04}{.02}) = P(-7 < Z < 3) = P(Z < 3) - P(Z < -7) = .9987$.

The second machine is more likely to produce acceptable corks.

39. The defect length of a corrosion defect in a pressurized steel pipe is normally distributed with mean value 30 mm and standard deviation 7.8 mm [suggested in the article **"Reliability Evaluation of Corroding Pipelines Considering Multiple Failure Modes and Time-Dependent Internal Pressure"** (***J. of Infrastructure Systems*, 2011: 216–224**)].

a. What is the probability that defect length is at most 20 mm? Less than 20 mm?

b. What is the 75th percentile of the defect length distribution—that is, the value that separates the smallest 75% of all lengths from the largest 25%?

c. What is the 15th percentile of the defect length distribution?

d. What values separate the middle 80% of the defect length distribution from the smallest 10% and the largest 10%?

Page 168, Problem 39

$\mu = 30$ mm and $\sigma = 7.8$ mm.

(a) $P(X \le 20) = P(Z \le -1.28) = .1003$. $P(X < 20) = P(X$

(b) Set $\Phi(z) = .75$, then $z = 0.67$. Let $(X - 30)/7.8 = .67$, th

(c) Set $\Phi(z) = .15$, then $z = -1.04$. Let $(X - 30)/7.8 = -1.0$

(d) The values in question are the 10th and 90th percentiles o

and $\Phi(z) = .9$ gives $z = 1.28$, thus the 10th and 90th perc

and 39.98 mm.

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