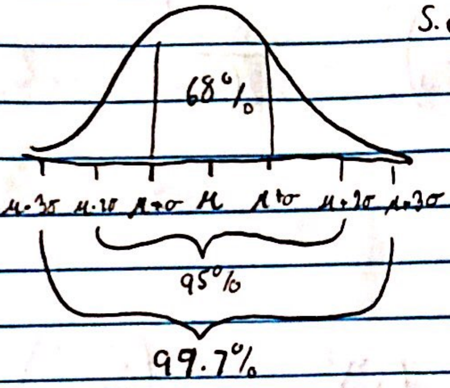
4/2/18

Empirical Rule

Mean = μ

S.d = σ



Example: College A and B both use SAT score for admission criteria. $\mu = 500$ $\sigma = 80$

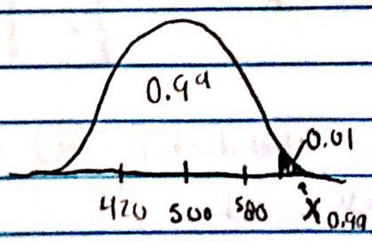
College A accepts people whose score is above 600.

Q1) What % of high school seniors can get into college A?

Score is X $\rightarrow$ $P(X > 600) = P\left(\frac{X-500}{80} > \frac{600-500}{80}\right)$

$= P\left(Z > \frac{5}{4}\right) = 1 - P(Z \leq 1.25) = 0.1056 = 10.56\%$

Q2) College B accepts the top 1%. What is the minimum score required to get accepted by college B?



$X = 500 + 80Z$

$P(X \leq X_{0.99}) = 0.99$

$P(500 + 80Z < X_{0.99}) = 0.99$

$P\left(Z \leq \frac{X_{0.99} - 500}{80}\right) = 0.99$

$Z_{0.99} = \frac{X_{0.99} - 500}{80}$ $\boxed{X_{0.99} = 500 + 80Z_{0.99}}$ $\boxed{= 686}$

□ Other common continuous distributions

Uniform

$a<b$ $\quad f(x) = \frac{1}{b-a} \quad x \in [a,b]$

$$E(x) = \int_a^b x \frac{1}{b-a} dx = \left(\frac{x^2}{2(b-a)}\right)_a^b = \frac{b^2-a^2}{2(b-a)}$$

$$= \frac{\cancel{(b-a)}(b+a)}{2\cancel{(b-a)}} = \frac{b+a}{2}$$

$$E[x^2] = \int_a^b x^2 \frac{1}{b-a} da = \frac{1}{b-a}\left(\frac{x^3}{3}\right)_a^b = \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right)$$

$$E(x^2)-E(x)^2 = \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{4b^3-4a^3-3(b+a)(b-a)}{12(b-a)}$$

$$= \frac{4b^3-4a^3-3b^3+3a^2b-3ab^2+3a^3}{12(b-a)}$$

varience

$$= \frac{b^3-a^3+3a^2b-3ab^2}{12(b-a)} = \frac{(b-a)^3}{12(b-a)} = \boxed{\frac{(b-a)^2}{12}}$$

Exponential $\lambda > 0$

$f(x) = \lambda e^{-\lambda x} \quad x > 0 \quad E(x) = \frac{1}{\lambda}$

$Var(x) = \frac{1}{\lambda^2}$

Gamma distribution $\alpha > 0, \ \beta > 0$

$F(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{\beta x} \quad E(x) = \frac{\alpha}{\beta}, \ V(x) = \frac{\alpha}{\beta^2}$

gamma

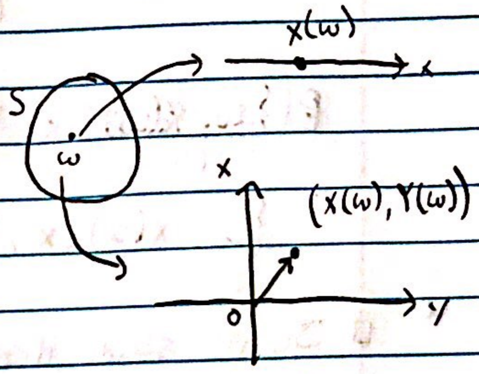
↓

Scanned by CamScanner

Sum of α independent identical exponential (λ) becomes gamma(α, λ)

Jointly distributed random variables

Random vector $\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$



A random vector (X, Y) is a vector valued function from S to $\mathbb{R}^2$

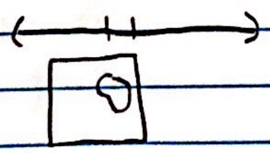
More generally, k-component random vector $(X_1 \dots X_k)$ is a function from S to $\mathbb{R}^k$, $k \geq 2$

Every component of a random vector is a random variable denote the range of (X, Y) to be $S_{X,Y}$

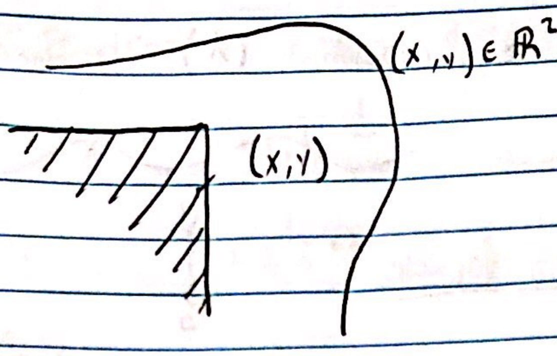
S_X: range of X

S_Y: range of Y

Joint Distribution Function of random variables



The (joint) distribution function of (X, Y) is a function from $\mathbb{R}^2$ to $\mathbb{R}^1$ is defined to be $\underbrace{F_{X,Y}(x,y) = P(X \leq x, Y \leq y)}_{\text{bivariant function}}$



$$P(\{\omega : X(\omega) \leq x,\ Y(\omega) \leq y\})$$

$$\{\omega : X(\omega) \leq x,\ Y(\omega) \leq y\}$$

- Joint: study distribution of (X,Y)
- Marginal: we study distributions of components of the random vector
  - ↳ study X
  - ↳ study Y
- If each component of (X,Y) is discrete we have joint pmf
- If each component of (X,Y) is continuous we have joint pdf

We need to extend domain of p and f accordingly

$\mathbb{Z}^2$ or $\mathbb{R}^2$