Kom505E

Week 3

04.10.2016

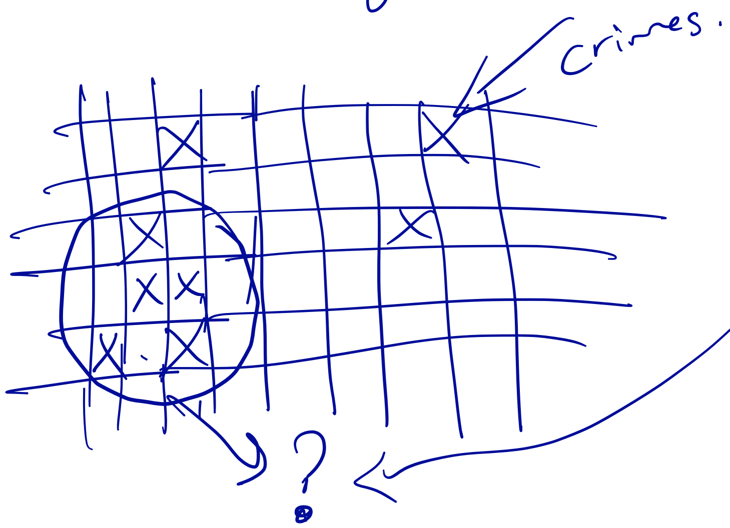
G.U.

Ch.4 Conditional Prob. + Bayes Thm.

4.7 Real World Ex:

$A = \{\text{observed crime data}\}$

$B = \{$ there is a cluster (gang) in the given data (shaded area) $\}$



$P(B|A)$ = posterior prob.

How can we make a decision to accept or reject this hypothesis?

$$\underline{P(B|A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

<u>Odds Ratio</u>: Another quantity: Odds against the hypothesis:

posterior prob. no cluster

$$\text{odds ratio} \triangleq \frac{P(B^c|A)}{P(B|A)} = \frac{P(A|B^c)P(B^c)/\cancel{P(A)}}{P(A|B)P(B)/\cancel{P(A)}}$$

posterior prob. cluster

$P(A|B^c) = ?$

Binomial law:

$$P(k \text{ crimes in } M \text{ cells}) = \binom{M}{k} p_{nc}^k (1-p_{nc})^{M-k}$$

$$P(A|B^c) = \binom{M}{k} P_{nc}^{k} (1-P_{nc})^{M-k}$$

$$= P(k=11 \mid \text{no cluster (gap)})$$

$$= \binom{145}{11} (0.01)^{11} (0.99)^{134}$$ 11 hits

$P_{nc} = \frac{29}{2500} \simeq 0.01$ from the data

M=145 region

50x50 area

$$P(A|B) = \binom{M}{k} P_c^{k} (1-P_c)^{M-k}$$

$P_c = \frac{11}{145}$ crimes / cells

prob. of a cluster (gap)

$P(B) = ?$ 10^{-6}: your belief (prior) that a cluster (gap) exists.

Due to your belief that a cluster is highly unlikely.

$$\text{Odds Ratio} = \frac{P(B^c|A)}{P(B|A)} = \frac{P(A|B^c)P(B^c)}{P(A|B)P(B)} = \frac{(0.01)^{11}(0.99)^{134}(1-10^{-6})}{(0.1)^{11}(0.9)^{134}(10^{-6})}$$

$$\simeq 3.52$$

$\Longrightarrow$ We can reject the hypothesis that a gap exists b/c the odds against it are 3.5 times higher.

Independent events:

$$P(ABC) = P(A)P(B)P(C)$$

When they are not independent, what happens?

For non-independent events:

$$P(ABC) = P(A|BC)\,\underbrace{P(BC)}$$

$$= P(A|BC)\cdot P(B|C)\,P(C)$$

↳ generalize

Probability Chain Rule:

$$P(A_1,\ldots A_N) = P(A_N|A_{N-1}\ldots A_1)\,P(A_{N-1}|A_{N-2}\ldots A_1)\ldots$$

$$\ldots P(A_2|A_1)\,P(A_1)$$

$\boxed{\text{Bernouilli trials}}$: 1 coin toss

$S_1 = \{H, T\}$

$p(H) = p$

$p(T) = 1-p$.

} 1 Bernouilli trial

~ coin toss.

2 coin tosses: $S_2 = S_1 \times S_1$

M coin tosses: $S_M = S_1 \times S_2 \times \ldots \times S_M$.

M independent trials.

P(getting k successes in M Bernouilli trials) $= p^k (1-p)^{M-k}$

k heads

$M-k$ tails

$\underbrace{HHH\ldots H}_{k}\;\underbrace{TT\ldots T}_{M-k}$

$\binom{M}{k}$

Dependent Bernoulli trials:

ex. (4.10) Two coins: 1 fair ($p=0.5$), 1 unfair coin ($p=0.25$)

Rule of experiment: i) Choose a coin at random.

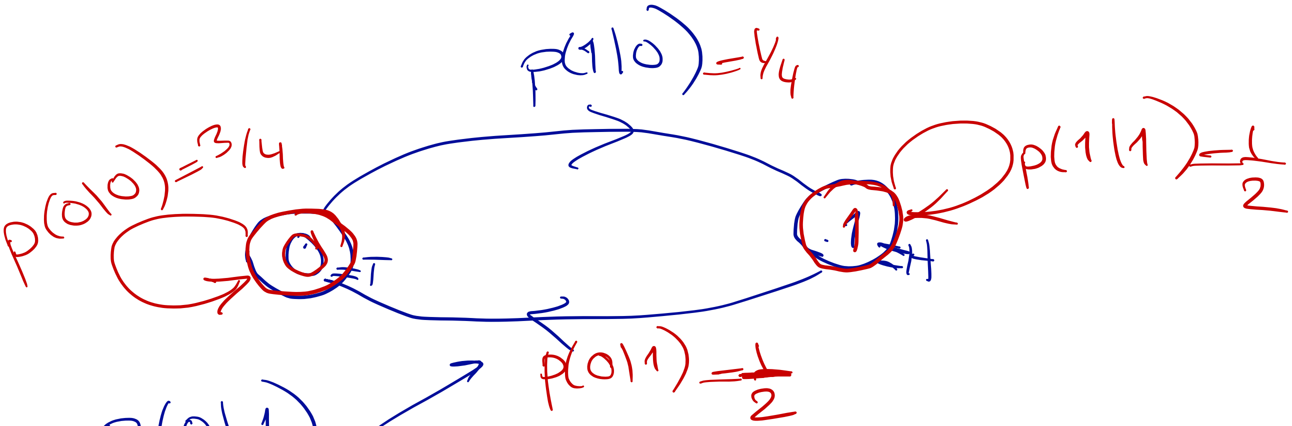
ii) Get a tail $\Rightarrow$ switch to unfair coin

Get a head $\Rightarrow$ switch to fair coin.

eg. $P(10 \text{ tails in succession}) = ?$

$\Rightarrow$ Clearly, the trials are not independent.

(Note: if we had both fair coins, indep. experiment.
10 coin toss $P(\cdot) = (0.5)^{10}$.)



$P(0|1)$

$P(1|0)$

$P(\text{tail on } i^{th} \text{ toss} \mid \text{tail on } (i-1)^{st} \text{ toss})$

$= P(A_i \mid A_{i-1}) \Rightarrow$ state transition prob.

T T . . . . T

0 0 . . . . 0 0 — A_i

10 tails.

Markov property.

$$P(A_1 A_2 \ldots A_{10}) = P(A_{10}|A_9)\,P(A_9|A_8) \ldots P(A_2|A_1)P(A_1)$$

$$P(A_1 \ldots A_{10}) = P(A_{10}|A_9 \ldots A_1)\,P(A_9|A_8 \ldots A_1) \ldots P(A_2|A_1)P(A_1)$$

$$\underbrace{P(A)}_{10\text{ tails}} = \prod_{i=2}^{10} \underbrace{P(A_i \mid A_{i-1})}_{P(0 \mid 0) = \frac{3}{4}} \, P(A_1) = \left(\frac{3}{4}\right)^9 \cdot \frac{5}{8} = 0{,}0469$$

$$P(A_1) = \underbrace{\frac{1}{2} \cdot \frac{3}{4}}_{\text{unfair}} + \underbrace{\frac{1}{2} \cdot \frac{1}{2}}_{\text{fair}} = \frac{5}{8}$$

$\frac{3}{4} \to p(\text{tail})$, $\frac{1}{2} \to p(\text{tail})$

$$\underbrace{P(A_1)}_{p(\text{tail})} = P(\text{tail} \mid \underbrace{\text{unfair}}_{B}) \, p(\underbrace{\text{selecting unfair}}_{B}) + p(\text{tail} \mid \underbrace{\text{fair}}_{B^c}) \, p(\underbrace{\text{select fair coin}}_{B^c})$$

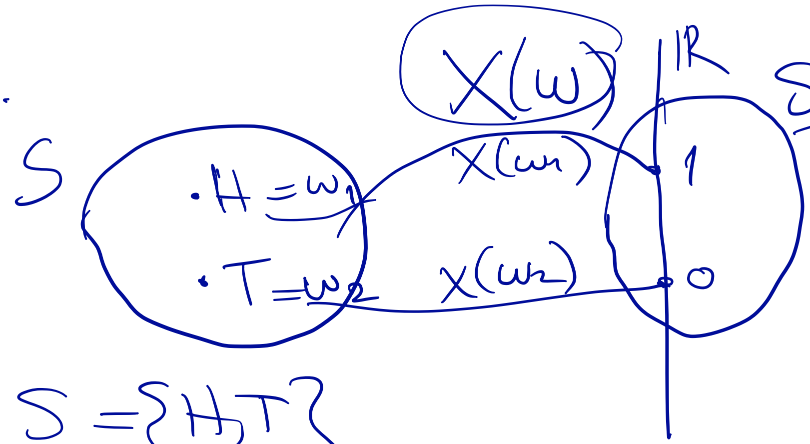
$$\Rightarrow \left(\frac{1}{2}\right)^{10} \longrightarrow \text{compare.}$$

Chapter 5 Discrete Random Variables :

From axioms, we know $\exists$ a sample space S, there are events & probabilities of events.

$\rightarrow$ Random variables (R.V.s) : derived quantity.

eg.



$S^X = \{0, 1\}$

$S = \{H, T\}$

Def : (RV) An RV X is a mapping (function) from sample space S to a subset of $\mathbb{R}$ (real line)

$\{x : -\infty \le x \le \infty\}$

RV ex:
$$X(\omega_i) = \begin{cases} 1, & \omega_1 = \text{head} \\ 0, & \omega_2 = \text{tail} \end{cases}$$

$$S^X = \{1, 0\} = \{X(\omega_1), X(\omega_2)\} = (X_1, X_2)$$

* In a discrete RV, X takes discrete values ; finite or countable {set of points $\in S^X$}

Discrete RV

$S^X = [0 : 0.01 : 1]$;

$|S^X| = 101$

Continuous RV

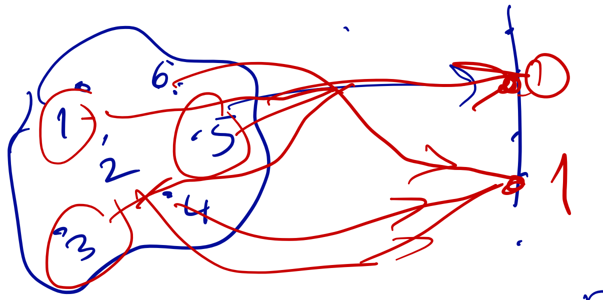
$S^X = [0, 1]$

$S^X = (-\infty, \infty)$

$|S^X| = \infty$

$\star$ $X(\cdot)$ could be a 1-1 or many-to-one mapping

eg: roll a die



$$X(w) = \begin{cases} 0, & w_i = 1,3,5 \\ 1, & w_i = 2,4,6 \end{cases}$$

Probability of an RV: $P[X(w) = x_i]$, $x_i \in S^x$.

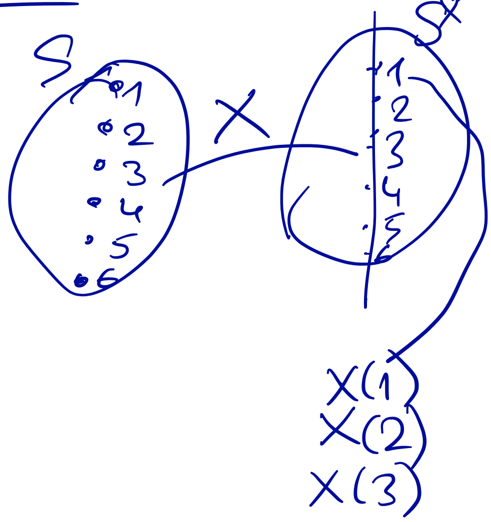
$$P_x[k] = P(X(w) = k)$$

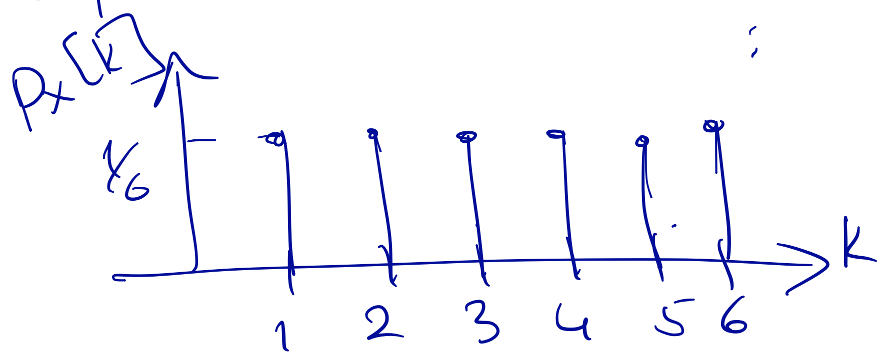
Def: Probability Mass Function (pmf) $P_x[k] = P(X = k)$

ex: Roll a fair die experiment: $X \Rightarrow \#$ on the die.

$P[X = 1] = \frac{1}{6}$

⋮





Properties of Pmf:

(1) $0 \leq P_x(k) \leq 1$.

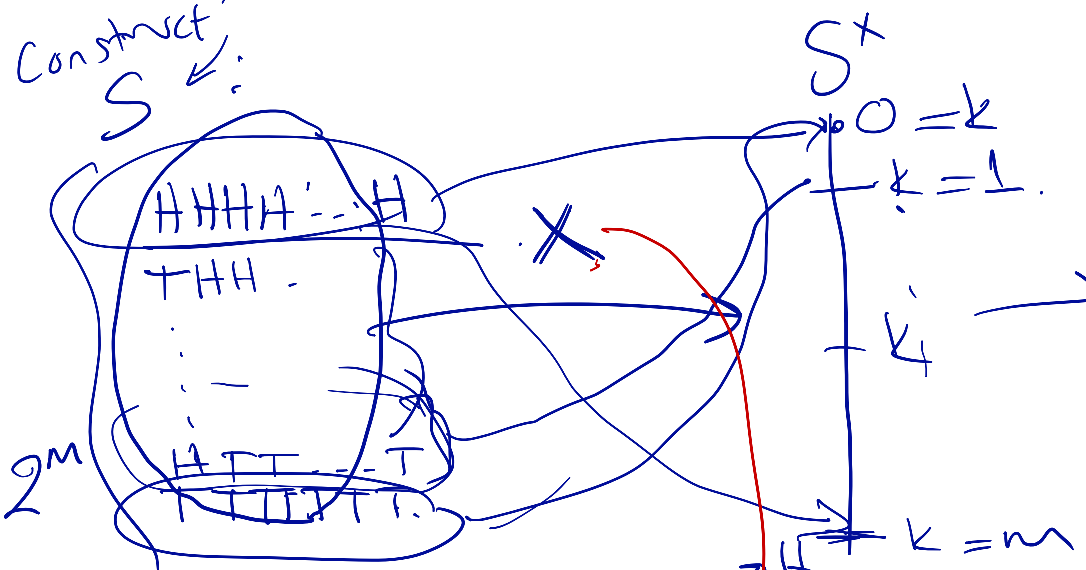
(2) $$\sum_{\forall x_i \in S^x} p[x_i] = 1$$

3) $$p(x \in A) = \sum_{\{i : \forall x_i \in A\}} p[x_i]$$

Ex: Binomial R.V. (r.v.) · $\begin{pmatrix}\text{M Bernoulli} \\ \text{trials}\end{pmatrix} \rightarrow \boxed{\frac{\text{k successes}}{\text{in M trials}}}$

$P_X[k] = P(X(\omega)=k) =$

Construct S:



$\rightarrow P(X=k)$

$= \binom{M}{k} p^k (1-p)^{M-k}$

$k = 0, 1, \ldots, M$

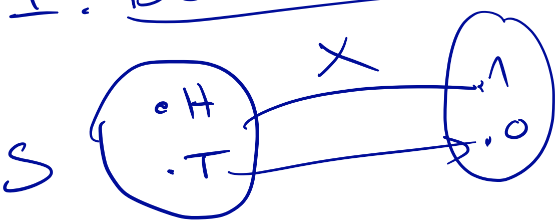
M trials : Bernoulli trials $\rightarrow$ H, T

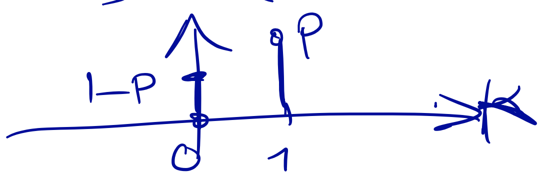
X maps # successes (Heads) to that # k

Important pmfs:

1. Bernoulli pmf:

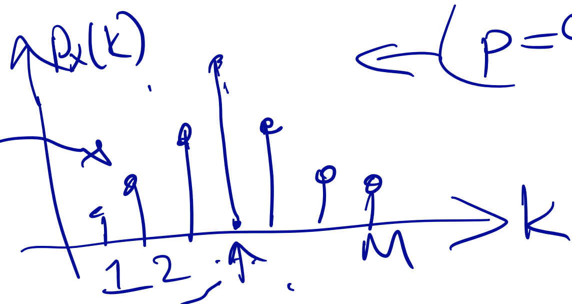
$p_X(k) = \begin{cases} p, & k=1 \\ (1-p), & k=0 \end{cases}$





2. Binomial pmf:

$p_X(k) = \binom{M}{k} p^k (1-p)^k$

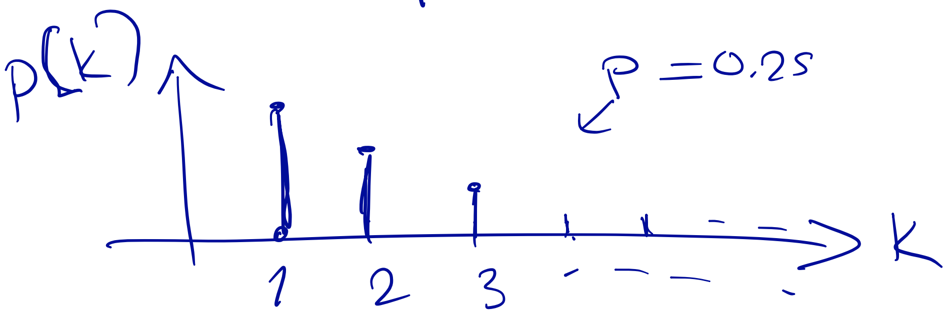


$\Leftarrow (p = 0.5)$

3) Geometric pmf : 1st success at k^{th} trial.
(build / based on Bernoulli)

$$\underbrace{T\,T\,\ldots\,T}_{k-1}\ \underset{k}{H} \rightarrow p(X=k) = (1-p)^{k-1}\cdot p$$

$k = 1, 2, \ldots \infty$



Check your book for the plots.
(exercise: construct S. .)

4) Poisson pmf : $p_X[k] = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $\lambda > 0$, $k > 0$

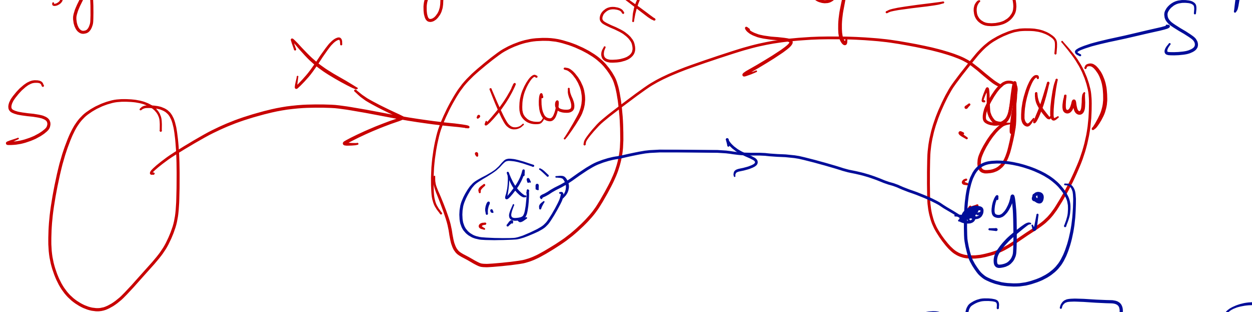
X : # arrival requests (e.g. no of people coming to a cashier in a given time.

λ : expected # arrivals in a unit time.

Ch. 5. Servicing customers : → Reading

Assignment: Read 5.10

Transformation of Discrete RVs :

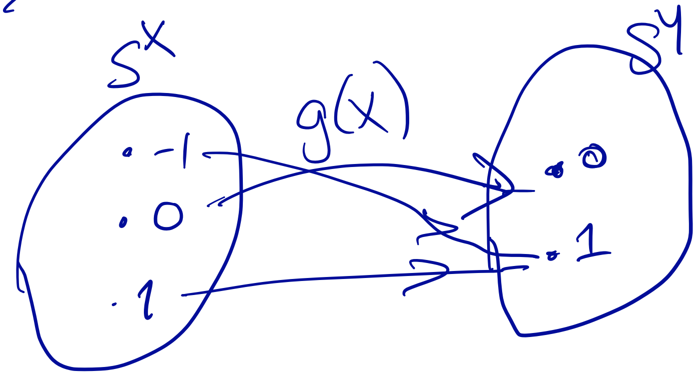


$P[y_i] = ?$

$$P_Y[y_i] = \sum_{\{j:\, g(x_j) = y_i\}} p_X[x_j]$$

Ex: $Y = X^2 = g(x) \Rightarrow S^X = \{-1, 0, 1\}$

$S^Y = \{0, 1\}$.



$g(x_j) = x_j^2 = 0$, for $x_j = 0$

$p_Y(y=0) = p_X[x=0]$

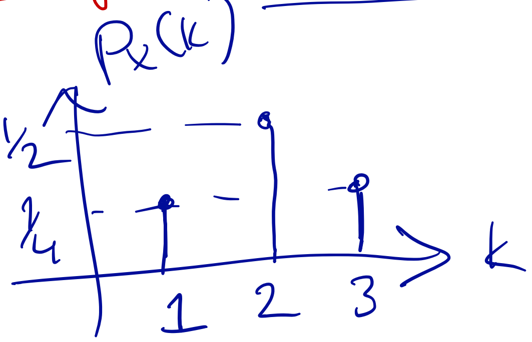
$p_Y[y=1] = p_X[x=-1] + p_X[x=1]$

$$p_X[k] = \begin{cases} 1/4, & k=-1 \\ 1/2, & k=0 \\ 1/4, & k=1 \end{cases} \Rightarrow p_Y[k] = \Big\{$$

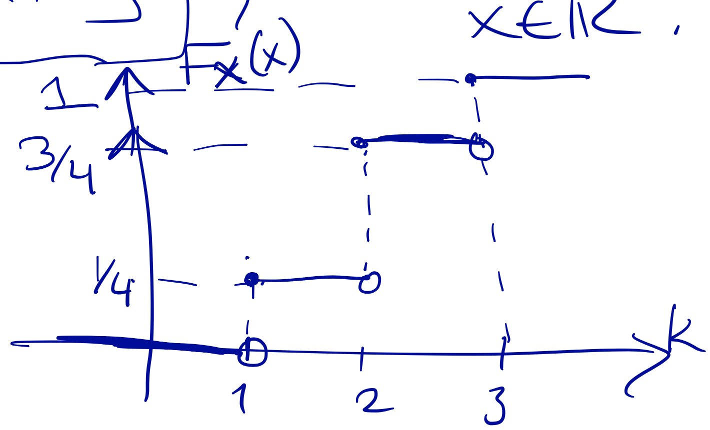
Exercise: Solve Ex 5.6 from your textbook; transform Poisson pmf.

Cumulative Distrib. Function (CDF).

Def: $\boxed{F_X(x) = P[X \le x]}$, $-\infty < x < \infty$, $x \in \mathbb{R}$.



$\Longrightarrow$



Read Properties of CDF Sec 5.8 in your textbook.