Distribution Reconstruction

COMP 3/4704: Data Privacy

1 Background Knowledge

- A person's salary is between $0 and $10,000 (integer)
- 1% of the population has no earning (zero salary)
  - Pr(S=0) = 1% = 0.01 (S is salary)
- Every other salary number is equally likely
  - Pr(S=k) = (1 - 0.01) / 10000 = 0.000099
- What is the chance that Bob has no earning?

2 Privacy Risk in Simple Perturbation

- Say a salary value S=k is changed as follows
  - R_1 : k stays the same with 20% chance, or changed to another number with 80% chance
- The method reports that Bob's salary is zero
- What is the chance that Bob has no earning?
  - previously, it was 1%
  - **it is now 96.1%**

Perturbed Data

- Consider the numeric attribute X
  - a random variable with values coming according to some probability density function f_X
- Say values of the attribute X have been modified by adding numbers from another random variable Y with density function f_Y
  - we call the new variable W
  - W = X + Y
- How can we estimate f_X using f_{X+Y} and f_Y?

Using Bayes Rule

say one value of W is w_i (= x_i+y_i)

estimated CDF

$$\Pr(X < t) = F_X^{'}(t) \equiv \int_{-\infty}^{t} f_X(z \mid X + Y = w_i)dz$$

$$= \int_{-\infty}^{t} \frac{f_{X+Y}(w_i \mid X = z) f_X(z)}{f_{X+Y}(w_i)} dz$$

$$= \int_{-\infty}^{t} \frac{f_{X+Y}(w_i \mid X = z) f_X(z)}{\int_{-\infty}^{\infty} f_{X+Y}(w_i \mid X = h) f_X(h)dh} dz$$

$$= \frac{\int_{-\infty}^{t} f_{X+Y}(w_i \mid X = z) f_X(z)dz}{\int_{-\infty}^{\infty} f_{X+Y}(w_i \mid X = z) f_X(z)dz}$$

$$= \frac{\int_{-\infty}^{t} f_Y(w_i - z) f_X(z)dz}{\int_{-\infty}^{\infty} f_Y(w_i - z) f_X(z)dz}$$

since we have multiple observed values (say *n* of them), we take the average in the estimated CDF

$$F_X^{'}(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{\int_{-\infty}^{t} f_Y(w_i - z) f_X(z)dz}{\int_{-\infty}^{\infty} f_Y(w_i - z) f_X(z)dz}$$

Estimated density function

- The estimated probability density function is the derivative of the estimated cumulative distribution function

$$f_X^{'}(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{f_Y(w_i - t)f_X(t)}{\int_{-\infty}^{\infty} f_Y(w_i - z)f_X(z)dz}$$

- We are still stuck because to compute this we need f_Y and f_X both
  - we do know f_Y but not f_X

Reconstruction Algorithm

- Iterative algorithm
  1. assume f_X = uniform distribution
  2. compute

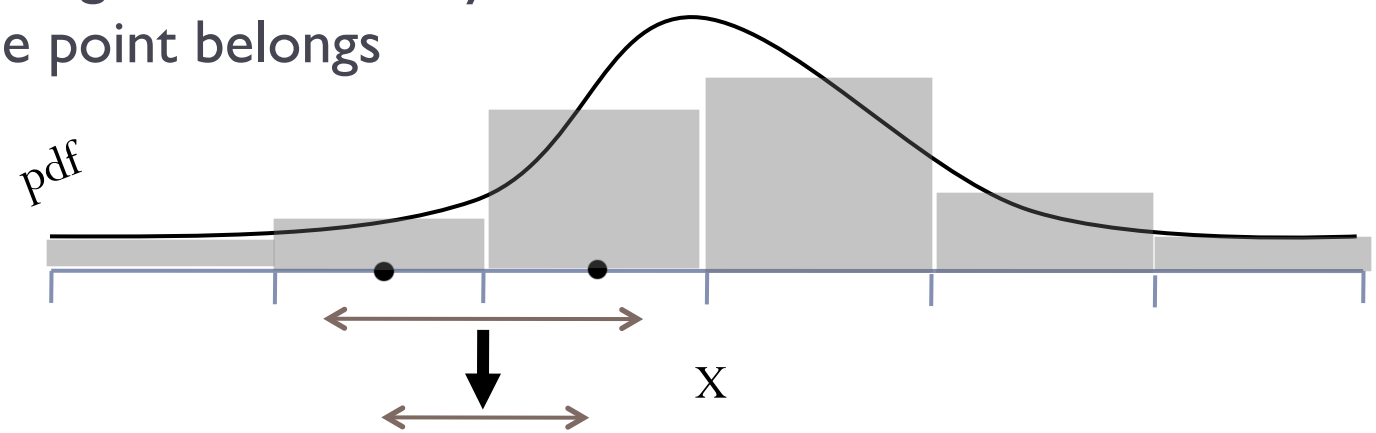
$$f_X(t) \leftarrow \frac{1}{n}\sum_{i=1}^{n}\frac{f_Y(w_i - t)f_X(t)}{\int_{-\infty}^{\infty} f_Y(w_i - z)f_X(z)dz}$$

  3. repeat (2) until difference in value is too small
     - e.g. 1% of threshold of some statistical test

- We cannot compute f_X for all possible values of t (infinite of them)
- To make step (2) simpler we will divide the domain of X into intervals

Using Intervals

- Divide the domain of X into intervals
  - $\boldsymbol{I_1}$: $[l_1,u_1]$, …, $\boldsymbol{I_m}$: $[l_m,u_m]$
- Let $mp_1, mp_2, \ldots, mp_m$ be the midpoints of these intervals
- We assume that
  - the distance between any two points is equal to the distance between the midpoints of the intervals they belong to
  - the value of the density function at a point is equal to the average of the density function values in the interval to which the point belongs



Using Intervals

$$f'_X(t) = \frac{1}{n}\sum_{i=1}^{n} \frac{f_Y(w_i - t) f_X(t)}{\int_{-\infty}^{\infty} f_Y(w_i - z) f_X(z) dz} \Rightarrow f'_X(t) = \frac{1}{n}\sum_{i=1}^{n} \frac{f_Y(m[w_i] - m[t]) g_X(I[t])}{\sum_{r=1}^{m} f_Y(m[w_i] - mp_r) g_X(I_r) L_r}$$

m[.] = *midpoint of the interval on which point lies*

I[.] = *interval on which point lies*

L_r = *length of r-th interval* $= u_r - l_r$

$g_X(I_r)$ = *average of* f_X *in an interval* I_r

$$= \int_{I_r} f_X(z) dz \Big/ L_r$$

estimation of **g**

$$g'_X(I_t) = \int_{I_t} f'_X(z) dz \Big/ L_t$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{f_Y(m[w_i] - mp_t) g_X(I_t)}{\sum_{r=1}^{m} f_Y(m[w_i] - mp_r) g_X(I_r) L_r}$$

Using Intervals

- Therefore we can estimate the probability for X being in an interval I_t

$$\Pr'(X \in I_t) = g'_X(I_t)L_t$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{f_Y(m[w_i] - mp_t)g_X(I_t)}{\sum_{r=1}^{m} f_Y(m[w_i] - mp_r)g_X(I_r)L_r} L_t$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{f_Y(m[w_i] - mp_t)\Pr(X \in I_t)}{\sum_{r=1}^{m} f_Y(m[w_i] - mp_r)\Pr(X \in I_r)}$$

- We use this in step (2) of the iterative algorithm

Reconstruction Algorithm

- You have intervals I_t (t=1…m) and want to estimate the probability that the attribute's value is in these intervals
- Iterative algorithm
  1. assume uniform distribution: $Pr_0(X \varepsilon I_t) = 1/m$
  2. let *iter*=1; compute (t=1…m)

$$\Pr_{iter}(X \in I_t) \leftarrow \frac{1}{n} \sum_{i=1}^{n} \frac{f_Y(m[w_i] - mp_t) \Pr_{iter-1}(X \in I_t)}{\sum_{r=1}^{m} f_Y(m[w_i] - mp_r) \Pr_{iter-1}(X \in I_r)}$$

  3. increment *iter* and repeat (2)
     - until difference in value is too small (1% of threshold of some statistical test), or
     - some maximum number of iterations