Probability Theory and Mathematical Statistics

Homework 2, Vitaliy Pozdnyakov

Part 1

1. Prove that the frequency of an event A is an efficient estimator of probability P(A)

Let $X_1, X_2, \dots, X_n$ — not dependent random variables, each distributed:

$P(A) = P(X_i = 1) = p$

$P(\overline{A}) = P(X_i = 0) = 1 - p = q$

Let frequency $\hat{\theta}_p = \frac{m}{n}$ where m — the number of happened events A (if $X_i = 1$) and n — the number of all attempts

$M(X_i) = 1 \cdot p + 0 \cdot q = p$

$M(X) = M(X_1 + X_2 + \dots + X_n) = nM(X_i) = np$

$M(\frac{m}{n}) = M(\frac{X_1+X_2+\dots+X_n}{n}) = \frac{1}{n}M(X) = \frac{np}{n} = p$

Thereby $\frac{m}{n}$ is an unbiased estimator

$D(X_i) = (1-p)^2 p + (0-p)^2 q = q^2 p + p^2 q = pq(p+q) = pq$

$D(X) = D(X_1 + X_2 + \dots + X_n) = npq$

$D(\frac{m}{n}) = D(\frac{X_1+X_2+\dots+X_n}{n}) = \frac{D(X)}{n^2} = \frac{pq}{n}$

Let $\varphi(1, p)$ and $\varphi(0, q)$ are PDFs of value 1 and 0 respectively, then

$\varphi(1, p) = p$ and $\varphi(0, p) = 1 - p$

Fisher information

$I(p) = M\left[\ln \varphi(X, p)'_p\right]^2 = \left[\frac{\varphi'(1,p)}{\varphi(1,p)}\right]^2 \varphi(1, p) + \left[\frac{\varphi'(0,p)}{\varphi(0,p)}\right]^2 \varphi(0, p) = \frac{[\varphi'(1,p)]^2}{\varphi(1,p)} + \frac{[\varphi'(0,p)]^2}{\varphi(0,p)} = \frac{1}{p}$
$+ \frac{1}{1-p} = \frac{1}{p(1-p)} = \frac{1}{pq}$

$\min D(\hat{\theta}_p) = \frac{1}{nI(p)} = \frac{pq}{n} = D(\frac{m}{n})$

Thereby $\frac{m}{n}$ is an efficent estimator

2. The sample set ($X_1 \dots X_n$) of random variable X corresponds the exponential distribution with parameter $\frac{1}{\theta}$ ($E(\frac{1}{\theta})$). Prove that $\overline{X}$ is an efficient estimator

Let $\lambda = \frac{1}{\theta}$ and $X_i \sim Exp(\lambda)$

$M(X_i) = \lambda^{-1}$

$M(\overline{X}) = M(\frac{X_1+\dots+X_n}{n}) = \frac{n\lambda^{-1}}{n} = \lambda^{-1}$

Thereby $\overline{X}$ is an unbiased estimator of parameter λ^{-1}

$D(X_i) = \lambda^{-2}$

$D(\overline{X}) = D(\frac{X_1+\dots+X_n}{n}) = \frac{n\lambda^{-2}}{n^2} = \frac{\lambda^{-2}}{n} = \frac{1}{n\lambda^2}$

Let $\varphi(x, \lambda) = \lambda e^{-\lambda x}$

Fisher information $I(\lambda) = M[\ln \varphi(x, \lambda)'_\lambda]^2 = M[(\ln \lambda - \lambda x)'_\lambda]^2 = M[\frac{1}{\lambda} - x]^2 = M[\frac{1}{\lambda^2} - \frac{2x}{\lambda} + x^2]$

$= \frac{1}{\lambda^2} - \frac{2}{\lambda}M[X] + M[X^2] = \frac{1}{\lambda^2} - \frac{2}{\lambda^2} + \frac{2!}{\lambda^2} = \lambda^{-2}$ and then

$I(\lambda^{-1}) = \lambda^2$

$\min D = \frac{1}{nI(\lambda^{-1})} = \frac{1}{n\lambda^2} = D(\overline{X})$

Thereby $\overline{X}$ is an efficent estimator of parameter $\lambda^{-1} = \theta$

3. The random variable X can be 0 or 1 with probability 0.5

- **a. What is mean μ and variance σ^2 for X?**

$\mu = M(X) = 0 \cdot 0.5 + 1 \cdot 0.5 = 0.5$

$\sigma^2 = D(X) = (0 - 0.5)^2 0.5 + (1 - 0.5)^2 0.5 = 0.25 \cdot 0.5 + 0.25 \cdot 0.5 = 0.25 = \frac{\mu}{2}$

b. The set of $X : X_1 \dots X_9$. Let's take a look at five different estimators of mean μ:

- $\hat{\mu}_1 = 0.45$
- $\hat{\mu}_2 = X_1$

- $\overset{\wedge}{\mu}_3 = \overline{X}$
- $\overset{\wedge}{\mu}_4 = X_1 + \frac{1}{3}X_2$
- $\overset{\wedge}{\mu}_5 = \frac{2}{3}X_1 + \frac{2}{3}X_2 - \frac{1}{3}X_3$

Which of estimators are unbiased? What is the bias for each estimator? Which one is the most efficient?

$M(\overset{\wedge}{\mu}_1) = M(0.45) = 0.45 \neq \mu \Rightarrow$ it is biased estimator, $Bias = 0.5 - 0.45 = 0.05$

$M(\overset{\wedge}{\mu}_2) = M(X_1) = \mu \Rightarrow$ it is unbiased estimator, $Bias = 0$

$M(\overset{\wedge}{\mu}_3) = M(\overline{X}) = \frac{1}{9}M(X_1 + \cdots + X_9) = \frac{0.5\cdot 9}{9} = 0.5 = \mu \Rightarrow$ it is unbiased estimator, $Bias = 0$

$M(\overset{\wedge}{\mu}_4) = M(X_1 + \frac{1}{3}X_2) = M(X_1) + \frac{1}{3}M(X_2) = \frac{15}{30} + \frac{5}{30} = \frac{20}{30} \neq \mu \Rightarrow$ it is biased estimator, $Bias = \frac{20}{30} - \frac{15}{30} = \frac{5}{30}$

$M(\overset{\wedge}{\mu}_5 = \frac{2}{3}M(X_1) + \frac{2}{3}M(X_2) - \frac{1}{3}M(X_3) = \frac{2}{3}\mu + \frac{2}{3}\mu - \frac{1}{3}\mu = \mu \Rightarrow$ it is unbiased estimator, $Bias = 0$

Let $\varphi(0, \mu) = \mu$ and $\varphi(1, \mu) = \mu$

Fisher information $I(\mu) = \left[\frac{\varphi'(1,\mu)}{\varphi(1,\mu)}\right]^2 \varphi(1, \mu) + \left[\frac{\varphi'(0,\mu)}{\varphi(0,\mu)}\right]^2 \varphi(0, \mu) = \frac{1^2}{\mu}\mu + \frac{1^2}{\mu}\mu = \frac{2}{\mu}$

$\min D = \frac{1}{nI(\mu)} = \frac{1}{9\frac{2}{\mu}} = \frac{\mu}{18}$

Consider only unbiased estimators

$D(\overset{\wedge}{\mu}_2) = D(X_1) = \frac{\mu}{2} \neq \min D \Rightarrow$ it is not efficent estimator

$D(\overset{\wedge}{\mu}_3) = D(\overline{X}) = \frac{1}{9^2}D(X_1 + \cdots + X_9) = \frac{9\mu}{9^2 2} = \frac{\mu}{18} = \min D$

Answer: $\overset{\wedge}{\mu}_3$ is an efficent estimator (the most efficent)

4. Topographer measures the square of a rectangle area. The measurement tool is set as the length X and width Y of area are independent. The real value for the length X is 10 and for the width Y is 5. The distribution of X and Y:

X	P(X)
8	1/4
10	1/4
11	1/2

Y	P(Y)
4	1/2
6	1/2

The estimation of the square $A = XY$ is a random variable.

- **Is X unbiased estimator of length? Prove**

$M(X) = 8\frac{1}{4} + 10\frac{1}{4} + 11\frac{1}{2} = \frac{8}{4} + \frac{10}{4} + \frac{22}{4} = \frac{40}{4} = 10 = X \Rightarrow$ it is unbiased estimator

- **Is Y unbiased estimator of width? Prove**

$M(Y) = 4\frac{1}{2} + 6\frac{1}{2} = 5 = Y \Rightarrow$ it is unbiased estimator

- **Is $A = XY$ unbiased estimator of square? Prove**

$XY = 50$

$M(XY) = \sum_{i=1}^{3} \sum_{j=1}^{2} x_i y_j P(X = x_i) P(Y = y_j)$

$= 8 \cdot 4 \cdot \frac{1}{8} + 8 \cdot 6 \cdot \frac{1}{8} + 10 \cdot 4 \cdot \frac{1}{8} + 10 \cdot 6 \cdot \frac{1}{8} + 11 \cdot 4 \cdot \frac{1}{4} + 11 \cdot 6 \cdot \frac{1}{4}$

$= \frac{32}{8} + \frac{48}{8} + \frac{40}{8} + \frac{60}{8} + \frac{88}{8} + \frac{132}{8}$

$= \frac{400}{8} = 50 = XY \Rightarrow$ it is unbiased estimator

5. The two measurements of a side of square are taken X_1, X_2. Let measurements be independent with mean α (α is a real length of the square) and σ^2. What is MSE of the area estimator $X_1 X_2$?

α^2 — the real area of square

$MSE(X_1 X_2) = D(X_1 X_2) + Bias(X_1 X_2, \alpha^2)^2$

- $D(X_1 X_2) = D(X_1)D(X_2) + \alpha^2 D(X_1) + \alpha^2 D(X_2) = \sigma^4 + \alpha^2\sigma^2 + \alpha^2\sigma^2 = \sigma^4 + 2\alpha^2\sigma^2$

- $M(X_1 X_2) = M(X_1)M(X_2) = \alpha^2 \Rightarrow Bias = 0$

$MSE(X_1 X_2) = \sigma^4 + 2\alpha^2\sigma^2$

6. The mean λ of Poisson distribution can be 1 or 2. The one measurement of random variable X has been taken. Two estimators of λ are defined as:

	X=0	X=1	X >= 2
$\hat{\lambda}$	1	1.5	2
$\tilde{\lambda}$	0	1	2

a) What is the mean and variance for $\hat{\lambda}$?

$$\text{Poisson distribution: } P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!}$$

$M(X) = \lambda$, $D(X) = \lambda$

For $\lambda = 1$

$M(\hat{\lambda}) = 1 \cdot P(X = 0, \lambda = 1) + 1.5 \cdot P(X = 1, \lambda = 1) + 2 \cdot P(X \geq 2, \lambda = 1)$

$= [P(X \geq 2, \lambda = 1) = 1 - P(X = 0, \lambda = 1) - P(X = 1, \lambda = 1)]$

$= e^{-1} + 1.5 \cdot e^{-1} + 2 \cdot (1 - 2e^{-1})$

$= 2.5e^{-1} + 2 - 4e^{-1} = 2 - 1.5e^{-1}$

In [6]:

```
import math
```

In [27]:

```
2 - 1.5 * math.exp(-1)
```

Out[27]:

```
1.4481808382428365
```

$M(\hat{\lambda}) \approx 1.45$

$D(\hat{\lambda}) = (1 - (2 - 1.5e^{-1}))^2 e^{-1} + (1.5 - (2 - 1.5e^{-1}))^2 e^{-1} + (2 - (2 - 1.5e^{-1}))^2 (1 - 2e^{-1})$

$= (1.5e^{-1} - 1)^2 e^{-1} + (1.5e^{-1} - 0.5)^2 e^{-1} + (1.5e^{-1})^2 (1 - 2e^{-1})$

In [16]:

```
((1.5 * math.exp(-1) - 1) ** 2) * math.exp(-1) \
    + ((1.5 * math.exp(-1) - 0.5) ** 2) * math.exp(-1) \
    + (1.5 * math.exp(-1) ** 2) * (1 - 2 * math.exp(-1))
```

Out[16]:

```
0.12852405430626074
```

$D(\hat{\lambda}) \approx 0.13$

For $\lambda = 2$

$M(\hat{\lambda}) = 1 \cdot P(X = 0, \lambda = 2) + 1.5 \cdot P(X = 1, \lambda = 2) + 2 \cdot P(X \geq 2, \lambda = 2)$

$= [P(X \geq 2, \lambda = 2) = 1 - P(X = 0, \lambda = 2) - P(X = 1, \lambda = 2)]$

$= e^{-1} + 1.5 \cdot 2e^{-2} + 2 \cdot (1 - e^{-1} - 2e^{-2})$

$= e^{-1} + 3e^{-2} + 2 - 2e^{-1} - 4e^{-2}$

$= 2 - e^{-1} - e^{-2}$

In [17]:

```
2 - math.exp(-1) - math.exp(-2)
```

Out[17]:

```
1.4967852755919449
```

$M(\hat{\lambda}) \approx 1.5$

$D(\hat{\lambda}) = (1 - (2 - e^{-1} - e^{-2}))^2 e^{-1} + (1.5 - (2 - e^{-1} - e^{-2}))^2 2e^{-2}$
$+ (2 - (2 - e^{-1} - e^{-2}))^2 (1 - e^{-1} - 2e^{-2})$

$= (-1 + e^{-1} + e^{-2})^2 e^{-1} + (-0.5 + e^{-1} + e^{-2})^2 2e^{-2} + (e^{-1} + e^{-2})^2 (1 - e^{-1} - 2e^{-2})$

In [19]:

```
(-1 + math.exp(-1) + math.exp(-2))**2 * math.exp(-1) \
    + (-0.5 + math.exp(-1) + math.exp(-2))**2 * 2 * math.exp(-2) \
    + (math.exp(-1) + math.exp(-2))**2 * (1 - math.exp(-1) - 2*math.exp(-2))
```

Out[19]:

```
0.18232202392867392
```

$D(\hat{\lambda}) \approx 0.18$

Answer:

for $\lambda = 1$: $M(\hat{\lambda}) \approx 1.45$, $D(\hat{\lambda}) \approx 0.13$

for $\lambda = 2$: $M(\hat{\lambda}) \approx 1.5$, $D(\hat{\lambda}) \approx 0.18$

b) What is the mean and variance for $\tilde{\lambda}$?

For $\lambda = 1$

$M(\tilde{\lambda}) = 0 \cdot P(X = 0, \lambda = 1) + 1 \cdot P(X = 1, \lambda = 1) + 2 \cdot P(X \geq 2, \lambda = 1)$

$= [P(X \geq 2, \lambda = 1) = 1 - P(X = 0, \lambda = 1) - P(X = 1, \lambda = 1)]$

$= e^{-1} + 2(1 - 2e^{-1})$

$= e^{-1} + 2 - 4e^{-1} = 2 - 3e^{-1}$

In [21]:

```
2 - 3 * math.exp(-1)
```

Out[21]:

```
0.896361676485673
```

$M(\tilde{\lambda}) \approx 0.9$

$D(\tilde{\lambda}) = (0 - (2 - 3e^{-1}))^2 e^{-1} + (1 - (2 - 3e^{-1}))^2 e^{-1} + (2 - (2 - 3e^{-1}))^2 (1 - 2e^{-1})$

$= (3e^{-1} - 2)^2 e^{-1} + (3e^{-1} - 1)^2 e^{-1} + 9e^{-2}(1 - 2e^{-1})$

In [34]:

```
(3 * math.exp(-1) - 2) ** 2 * math.exp(-1) \
    + (3 * math.exp(-1) - 1) ** 2 * math.exp(-1) \
    + 9 * math.exp(-2) * (1 - 2 * math.exp(-1))
```

Out[34]:

```
0.6213796567276975
```

$D(\tilde{\lambda}) \approx 0.62$

For $\lambda = 2$

$M(\tilde{\lambda}) = 0 \cdot P(X = 0, \lambda = 2) + 1 \cdot P(X = 1, \lambda = 2) + 2 \cdot P(X \geq 2, \lambda = 2)$

$= [P(X \geq 2, \lambda = 2) = 1 - P(X = 0, \lambda = 2) - P(X = 1, \lambda = 2)]$

$= 2e^{-2} + 2 \cdot (1 - e^{-1} - 2e^{-2})$

$= 2e^{-2} + 2 - 2e^{-1} - 4e^{-2}$

$= 2 - 2e^{-1} - 2e^{-2}$

In [35]:

```
2 - 2 * math.exp(-1) - 2 * math.exp(-2)
```

Out[35]:

```
0.9935705511838899
```

$M(\tilde{\lambda}) \approx 0.99$

$D(\tilde{\lambda}) = (0 - (2 - 2e^{-1} - 2e^{-2}))^2 e^{-1} + (1 - (2 - 2e^{-1} - 2e^{-2}))^2 2e^{-2}$
$+ (2 - (2 - 2e^{-1} - 2e^{-2}))^2 (1 - e^{-1} - 2e^{-2})$

$= (-2 + 2e^{-1} + 2e^{-2})^2 e^{-1} + (-1 + 2e^{-1} + 2e^{-2})^2 2e^{-2} + (2e^{-1} + 2e^{-2})^2 (1 - e^{-1} - 2e^{-2})$

In [36]:

```
(-2 + 2*math.exp(-1) + 2*math.exp(-2))**2 * math.exp(-1) \
    + (-1 + 2*math.exp(-1) + 2*math.exp(-2))**2 * 2 * math.exp(-2) \
    + (2*math.exp(-1) + 2*math.exp(-2))**2 * (1 - math.exp(-1) - 2*math.exp(-2))
```

Out[36]:

```
0.7292880957146957
```

$D(\tilde{\lambda}) \approx 0.73$

Answer:

for $\lambda = 1$: $M(\hat{\lambda}) \approx 0.9$, $D(\hat{\lambda}) \approx 0.62$

for $\lambda = 2$: $M(\hat{\lambda}) \approx 0.99$, $D(\hat{\lambda}) \approx 0.73$

c) Which estimator is more accurate?

For $\lambda = 1$

$MSE(\hat{\lambda}) = D(\hat{\lambda}) + Bias(\hat{\lambda}, \lambda)^2 \approx 0.13 + 0.45^2$

In [37]:

```
0.13 + 0.45 ** 2
```

Out[37]:

```
0.3325
```

≈ 0.33

$MSE(\tilde{\lambda}) = D(\tilde{\lambda}) + Bias(\tilde{\lambda}, \lambda)^2 \approx 0.62 + 0.1^2$

≈ 0.621

For $\lambda = 2$

$MSE(\hat{\lambda}) = D(\hat{\lambda}) + Bias(\hat{\lambda}, \lambda)^2 \approx 0.18 + 0.5^2$

In [38]:

```
0.18 + 0.5 ** 2
```

Out[38]:

```
0.43
```

≈ 0.43

$MSE(\tilde{\lambda}) = D(\tilde{\lambda}) + Bias(\tilde{\lambda}, \lambda)^2 \approx 0.73 + 0.01^2$

≈ 0.73

Answer: $\hat{\lambda}$ is more accurate then $\tilde{\lambda}$ in both cases

7. Let X_1,X_2,X_3 be a set of values for random variable with mean (μ) and variance (σ^2). Let's take a look at two estimators of σ^2:

1. $\overset{\wedge}{\sigma_1^2} = c_1(X_1 - \mu)^2$
2. $\overset{\wedge}{\sigma_2^2} = c_2(X_1 - \mu)^2 + c_2(X_2 - \mu)^2 + c_2(X_3 - \mu)^2$

a) What c_1 ,c_2 can be equal to satisfy the $\overset{\wedge}{\sigma_1^2}$, $\overset{\wedge}{\sigma_2^2}$ are unbiased estimators of σ^2?

$$M(\overset{\wedge}{\sigma_1^2}) = M[c_1(X_1 - \mu)^2] = c_1 M[(X_1 - \mu)^2] = c_1\sigma^2 \Rightarrow c_1 = 1$$

$$M(\overset{\wedge}{\sigma_2^2}) = M[c_2(X_1 - \mu)^2 + c_2(X_2 - \mu)^2 + c_2(X_3 - \mu)^2] = c_2 M[(X_1 - \mu)^2]$$
$$+ c_2 M[(X_2 - \mu)^2] + c_2 M[(X_3 - \mu)^2)] = 3c_2\sigma^2 \Rightarrow c_2 = \frac{1}{3}$$

b) Find proportion: efficiency($\overset{\wedge}{\sigma_1^2}$) / efficiency($\overset{\wedge}{\sigma_2^2}$)

If $c_1 = 1$ and $c_2 = \frac{1}{3}$ then $Bias = 0$, so

$$MSE(\overset{\wedge}{\sigma_1^2}) = D(\overset{\wedge}{\sigma_1^2}) = D[(X_1 - \mu)^2]$$

$$MSE(\overset{\wedge}{\sigma_2^2}) = D[\tfrac{1}{3}(X_1 - \mu)^2 + \tfrac{1}{3}(X_2 - \mu)^2 + \tfrac{1}{3}(X_3 - \mu)^2] = \tfrac{1}{9} D[(X_1 - \mu)^2]$$
$$+ \tfrac{1}{9} D[(X_2 - \mu)^2] + \tfrac{1}{9} D[(X_3 - \mu)^2] = \tfrac{1}{3} D[(X_i - \mu)^2]$$

Thereby $\frac{MSE(\overset{\wedge}{\sigma_1^2})}{MSE(\overset{\wedge}{\sigma_2^2})} = \frac{D[(X_1-\mu)^2]}{\frac{1}{3}D[(X_i-\mu)^2]} = 3$

Part 2

1. (The mean of the empirical distribution)^2 is an estimator of (the mean of the random variable)^2 ? Proof

Let $\overline{X}$ be the mean of the emperical distribution and μ is the mean of the random variable

$$D(\overline{X}) = M(\overline{X}^2) - M(\overline{X})^2 \Rightarrow M(\overline{X}^2) = D(\overline{X}) + M(\overline{X})^2 = D(\overline{X}) + \mu^2$$

$D(\overline{X}) = D(\frac{X_1+\cdots+X_n}{n}) = \frac{1}{n^2}D(X_1 + \cdots + X_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$

Thereby $\overline{X}^2$ is biased estimator of μ^2 with $Bias = \frac{\sigma^2}{n}$

2. Let X be a random variable $\sim N(147.8, 12.3^2)$.

a. What is $P(X < 163.3)$?

$a = 147.8, \sigma^2 = 12.3^2$

$$P(X < 163.3) = \int_{-\infty}^{163} \frac{1}{\sqrt{2\pi \cdot 12.3^2}} e^{-\frac{(x-147.8)^2}{2 \cdot 12.3^2}} dx \approx 0.89$$

b. Let $\overline{X}$ be an empirical mean and s^2 be an empirical variance that have been obtained via sample set with size $n = 25$ of random variable X. What is $P(\overline{X} \leq 150.9)$?

$\overline{X} \leq 150.9 \Rightarrow \frac{X_1+\cdots+X_{25}}{25} \leq 150.9 \Rightarrow X_1 + \cdots + X_n \leq 3772.5$

Let $Y = X_1 + \cdots + X_{25} \Rightarrow M(Y) = M(X_1 + \cdots + X_{25}) = 25 \cdot 147.8 = 3695$

and $D(Y) = D(X_1 + \cdots + X_{25}) = 25 \cdot 12.3^2 = 3782.25$

Thereby $Y \sim N(3695, 3782.25)$

$$P(Y < 3772.5) = \int_{-\infty}^{3772.5} \frac{1}{\sqrt{2\pi \cdot 3782.25}} e^{-\frac{(x-3695)^2}{2 \cdot 3782.25}} dx \approx 0.83$$

c. What a and b can be equal to satisfy the $P(a \leq s^2 \leq b) = 0.95$?

$s^2 = \frac{(X_1-147.8)^2+\cdots+(X_{25}-147.8)^2}{25}$

Let $Y = X - 147.8 \Rightarrow M(Y) = M(X - 147.8) = M(X) - 147.8 = 0$

and $D(Y) = D(X - 147.8) = D(X) - 0 = 12.3^2$

Thereby $Y \sim N(0, 12.3^2) \Rightarrow Z = \frac{Y-0}{12.3} \sim N(0, 1)$

and $a \leq s^2 \leq b \Rightarrow a \leq \frac{12.3(Z_1^2+\cdots+Z_{25}^2)}{25} \leq b \Rightarrow \frac{25}{12.3}a \leq Z_1^2 + \cdots + Z_{25}^2 \leq \frac{25}{12.3}b$

$\Rightarrow \frac{25}{12.3}a \leq W \leq \frac{25}{12.3}b$ where $W \sim \chi^2(k = 25)$

Let $\frac{25}{12.3}a = 0 \Rightarrow P(0 \leq W \leq 37.652) \approx 0.95$

$37.652 = \frac{25b}{12.3} \Rightarrow b = \frac{37.652 \cdot 12.3}{25}$

In [65]:

```
37.652 * 12.3 / 25
```

Out[65]:

```
18.524784
```

Answer: $a = 0$, $b \approx 18.52$

3.The table shows the average temperature in January per year in two cities:

City \ Year	1891	1892	1893	1894	1895	1896	1897
City 1	-19.2	-14.8	-19.6	-11.1	-9.4	-16.9	-13.7
City 2	-21.8	-15.4	-20.8	-11.3	-11.6	-19.2	-13.0

City \ Year	1898	1899	1900	1901	1902	1903	
City 1	-4.9	-13.9	-9.4	-8.3	-7.9	-5.3	
City 2	-7.4	-15.1	-14.4	-11.1	-10.5	-7.2	

a) What is empirical mean?

$\overline{X} = \frac{1}{n} \sum x_i$

In [77]:

```
import pandas as pd
```

In [40]:

```
df = pd.DataFrame({
    'city1': [-19.2, -14.8, -19.6, -11.1, -9.4, -16.9, -13.7, -4.9, -13.9, -9.4, -8.
    'city2': [-21.8, -15.4, -20.8, -11.3, -11.6, -19.2, -13.0, -7.4, -15.1, -14.4, -
```

In [67]:

```
mean1 = 0
for i in list(df['city1']):
    mean1 += i
mean1 = mean1 / len(df)
print('The emperical mean for City 1:', round(mean1, 2))

mean2 = 0
for i in list(df['city2']):
    mean2 += i
mean2 = mean2 / len(df)
print('The emperical mean for City 2:', round(mean2, 2))
```

```
The emperical mean for City 1: -11.88
The emperical mean for City 2: -13.75
```

b) What is empirical variance?

$s^2 = \frac{1}{n} \sum (x_i - \overline{X})^2$ (a biased estimator)

In [71]:

```
var1 = 0
for i in list(df['city1']):
    var1 += (i - mean1) ** 2
var1 = var1 / len(df)
print('The emperical variance for City 1:', round(var1, 2))

var2 = 0
for i in list(df['city2']):
    var2 += (i - mean2) ** 2
var2 = var2 / len(df)
print('The emperical variance for City 2:', round(var2, 2))
```

```
The emperical variance for City 1: 22.14
The emperical variance for City 2: 20.09
```

c) What is empirical correlation coefficient?

$$corr = \frac{cov}{s_1 s_2} = \frac{\overline{X_1 X_2} - \overline{X_1} \cdot \overline{X_2}}{s_1 s_2}$$

In [75]:

```
prod_mean = 0
for i in range(len(df)):
    prod_mean += list(df['city1'])[i] * list(df['city2'])[i]
prod_mean = prod_mean / len(df)
cov = prod_mean - mean1 * mean2
corr = cov / (var1**(1/2) * var2**(1/2))
print('The emperical correlation coefficient:', round(corr, 2))
```

```
The emperical correlation coefficient: 0.96
```

d) Create the histogram using python/R

In [82]:

```
import matplotlib.pyplot as plt
```

In [98]:

```
df.hist(figsize=(10,5), bins=5, grid=False)
plt.show()
```