Written exam in Probability & Statistics

PM6 & ET6
Lecturer: Kasper K. Berthelsen

Tuesday 4th of January 2005, 9:00-13:00

In the assessment emphasis will be put on both correct methods as wells as correct answer, hence the the method should be clearly stated.

Good luck!

Problem 1. (approx. 20%)
A random variable X has mean 10 and variance 50.

1. Calculate the mean and variance of the random variable $Y = 10 + 5X$.
2. Find the mean of $V = (X - 10)^2$ and $Z = X^2$.

1) $E[Y] = E[10+5X] = 10 + 5E[X] = 10 + 5 \cdot 10 = \underline{\underline{60}}$

2) $Var[X] = E[(X - E[X])^2] = E[(X-10)^2] = \underline{\underline{50}}$

$Var[X] = E[X^2] - E[X]^2 = 50 \Rightarrow E[X^2] = 50 + E[X]^2 = 50 + 10^2 = \underline{\underline{150}}$

Problem 2. (approx. 10%)
In a survey among house owner, people are asked if they are willing to pay more for snow removal. Among the 84 people who reply the answers are distributed according to age as follow:

Age	No	Yes	No answer
20–25	1	0	0
26–35	0	3	1
36–50	6	10	10
51–60	1	7	1
61–70	2	13	6
> 70	4	13	6

1) $P(\{Yes\} \mid \{Age \leq 50\}) = \dfrac{P(\{Yes\} \cap \{Age \leq 50\})}{P(\{Age \leq 50\})} = \dfrac{13/84}{31/84} = \underline{\underline{0,42}}$

2) $P(\{Yes\} \mid \{Age > 50\}) = \dfrac{P(\{Yes\} \cap \{Age > 50\})}{P(\{Age > 50\})} = \dfrac{33/84}{53/84} = \underline{\underline{0,62}}$

2) Indept. if only if

$P(A \cap B) = P(A)\,P(B)$

$P(A \cap B) = \frac{13}{84} = 0,154$

$\left.\begin{array}{l} P(A) = \frac{31}{84} \\ P(B) = \frac{46}{84} \end{array}\right\} P(A)P(B) = 0,202$

Conclusion: <u>A and B dependt</u>.

1. Calculate the conditional probability for answering yes, conditionally on the person's age being ≤ 50 years and > 50 years, respectively.
2. Are the events $A = \{\text{Age} \leq 50 \text{ years}\}$ and $B = \{\text{Yes}\}$ independent? Justify your answer.

Problem 3. (approx. 20%)
In an airport, whenever the metal detector goes off, there is a 25% probability that the alarm is caused by coins in the pocket of the passenger walking through the metal detector.

1. During one day the alarm goes off 15 times. What is the probability that at least 3 of these alrams are caused by passengers having coins in their pockets?
2. Question 1 continued: Is it likely that none of these 15 alarms are caused by coins in a pocket? Explain your answer based on the probability of this event.
3. Just before Christmas the airport is unusually busy. On one day the metal detector alarm goes off 50 times. What is the probability that at most $\frac{1}{5}$ of these alarms are caused by coins in a pocket.

1) $X \sim B(15, 0,25)$ $\quad P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0,2361 = \underline{0,7639}$

2) $P(X=0) = 0,0134$, <u>fairly unlikely</u>

3) $X \sim B(50, 0,25)$, $E[X] = 50 \cdot 0,25 = 12,5$ $\quad V[X] = 0,25 \cdot (1-0,25) \cdot 50 = 9,375$

$X \sim N(12,5,\ 9,375)$ $\quad P(X \leq \frac{50}{5}) = P\left(\frac{Z - 12.5}{\sqrt{9.375}} \leq \frac{10 - 12.5}{\sqrt{9.375}}\right) = P(Z \leq -0,82) = \underline{\underline{0,2061}}$

1) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{4227}{14} = 301.93$ $\quad s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} = \frac{n\sum x_i^2 - (\sum x_i)^2}{n(n-1)} =$ $\quad s=\sqrt{s^2} = 69.66$

$= \frac{14\cdot 1339337 - (4227)^2}{14\cdot 13} = 4832,687$

2) $(1-\alpha)100\%$ CI: $\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = 301.93 \pm 2.160\cdot\frac{69.66}{\sqrt{14}} = [308.73, 295.13]$

Problem 4. (approx. 35%)
As is well-known, the department network is often down. Near the project dead-line some students decide to measure the dayly downtime in minutes. Acordingly they measure how many minutes the network is down each day for 14 days and obtain the following downtimes:

Day	1	2	3	4	5	6	7
downtime (minutes)	229	295	343	337	282	313	262
Day	8	9	10	11	12	13	14
downtime (minutes)	303	201	374	376	343	406	163

3) $(1-\alpha)100\%$ CI for σ^2

$\left[\frac{(n-1)s^2}{\chi_{\alpha/2,n-1}} ; \frac{(n-1)s^2}{\chi_{1-\alpha/2,n-1}}\right] = \left[\frac{13\cdot 4832,687}{24,74} ; \frac{13\cdot 4832,687}{5,01}\right]$

$= [2539.86 ; 12543,035]$

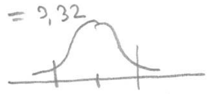
The downtimes are assumed to follow a normal distribution with mean μ and variance σ^2.

1. Estimate the mean μ and the standard deviation σ for the dayly downtime.
2. Determine a 95% confidence interval for μ.
3. Determine a 95% confidence interval for σ.
4. The students want the downtime to be as short as possible. Test on the 5% significance level if the expected downtime is significantly less than 4 hours, i.e. 240 minutes.
5. What is the probability that the average down time over a 14 day period is less then 4 hours, i.e. 240 minute? Assume that the 14 downtimes are independent and normal distributed with equal means $\mu = 300$ and unknown and equal variances.

4) $H_0: \mu \geq 240 \quad H_1: \mu < 240$

$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{301,93-240}{69,66/\sqrt{14}}$

$= 3,32$



$-t_{\alpha,n-1} = -2,16$

We cannot reject H_0

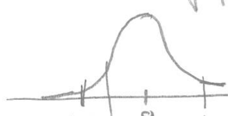
5) $P(\bar{x} < 240) = P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{240-\mu}{\sigma/\sqrt{n}}\right) = P\left(Z < \frac{-60\cdot\sqrt{14}}{\sigma}\right)$

Problem 5. (approx. 15%)
At a Christmas dinner 10 students measure their blood alcohol level by each making two measurements using a breathalyzer. They obtain the following measurements where the differences between measurements are given:

Student	1st measurement	2nd measurement	Difference
1	0.9	0.9	0.0
2	1.0	1.8	-0.8
3	1.8	1.8	0.0
4	1.2	1.6	-0.4
5	0.8	0.8	0.0
6	1.0	0.8	0.2
7	0.9	1.0	-0.1
8	1.2	2.1	-0.9
9	2.2	2.0	0.2
10	1.2	1.5	-0.3
$\bar{x}$	1.22	1.43	-0.21
s^2	0.197	0.260 (forkerte)	0.150

2) $H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 \neq \mu_2 \quad \alpha = 0,05$

$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{-0,21}{\sqrt{\frac{0,150}{10}}} = -1,71$



$-2,62 \quad 0 \quad -1,71 \quad t_{\alpha/2,n-1} = t_{0,025,9} = 2,62$

Cannot reject H_0

It is assumed that the random variables correspoindg to the alchohol level for first and second measurments are independent and noraml distributed with equal mean and variance. Notice that the two measurements for the same student are **not** independent.

1) $(1-\alpha)100\%$ CI: $\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

$-0,21 \pm 1,833\cdot\sqrt{\frac{0,150}{10}}$

$= [-0,435 ; 0,0145]$

1. Find a 90% confidence interval for the difference in the two measurements.
2. Test at the 5% significance level if the level at the first measurement is different from the second measurement.

Remember to add student number on all sheets and state how many sheets your solution consists of.

Written exam in Probability & Statistics

PM6 & ET6
Lecturer: Kasper K. Berthelsen

Friday 6th of January 2006, 9:00-13:00

In the assessment emphasis will be put on both correct methods as wells as correct answer, hence the the method should be clearly stated.

Good luck!

Problem 1. (approx. 20%)
A salesman at a used car dealer receives a commission for each car or van he sells. When he sells a car he receives 4200 kr and 4800 kr when he sells a van. He expects to sell a number of cars and vans each day according to the following probabilities:

X Number of cars	0	1	2	3
Probability	0.3	0.4	0.2	0.1

Y Number of vans	0	1	2
Probability	0.4	0.5	0.1

1. Calculate the expected number of cars and vans the salesman is expected to sell each day. $\mu_x = E[X] = \sum_x x\,P(x) = 1{,}1 \quad \mu_y = E[Y] = 0{,}7$
2. Calculate the standard deviation of the number of cars and van the salesman sells in a day. $\sigma_x^2 = \sum_x (x-\mu_x)^2 P(x) = \sum_x x^2 p(x) - E[X]^2 = 0{,}89 \quad \sigma_x = 0{,}94 \quad \sigma_y^2 = 0{,}41 \quad \sigma_y = 0{,}64$
3. Calculate the expected commission for both cars and vans a salesman will receive in a day. $E[4200X + 4800Y] = 4200\,E[X] + 4800\,E[Y] = 4200 \cdot 1{,}1 + 4800 \cdot 0{,}7 = 7980$
4. Calculate the standard deviation of the salesman total commission in a day when we assume that the number of sold cars and sold vans are dependent with a correlation coefficient of $\rho = 0.1$.

$\rho = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$

$Var(4200X + 4800Y) = 4200^2\,Var(X) + 4800\,Var(Y) + 2 \cdot 4200 \cdot 4800 \cdot Cov(X,Y)$

$= 4200^2 \cdot 0{,}89 + 4800^2 \cdot 0{,}41 + 2 \cdot 4200 \cdot 4800 \cdot 0{,}1 \cdot 0{,}94 \cdot 0{,}64$

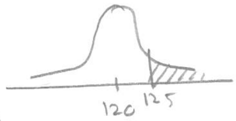
$Cov(X,Y) = \rho\,\sigma_x \sigma_y$

$= 27571165{,}1$

$\sqrt{\#} = 5250.87$

Problem 2. (approx. 15%)
The length of times it takes to repair a vending machine follows a normal distribution with mean 120 minutes and variance 16 minutes^2. If the vending machine is under repair for more than 125 minutes the machines must be cleaned and emptied which is an unwanted extra expense.

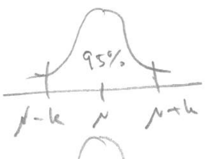
1. What is the probability that the vending machine is under repair for more than 125 minutes?
2. A member of staff wants to find a time interval in which the time it takes to repair the vending machine is with 95% probability. Find such a 95% probability interval which is symmetric around the mean.



$X \sim N(120, 16)$

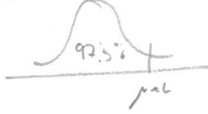
1) $P(X > 125) = P\left(\frac{X-120}{4} > \frac{125-120}{4}\right) = P(Z > 1{,}25) = 1 - P(Z < 1{,}25) = 1 - 0{,}89$
$= 0{,}11$

2) $P(\mu - k < X < \mu + k) = 0.95$



$P(X < \mu + k) = 0{,}975$

$P\left(\frac{X-\mu}{\sigma} < \frac{\mu+k-\mu}{\sigma}\right) = P\left(Z < \frac{k}{\sigma}\right) = 0{,}975$



$P(Z < 1{,}96) = 0{,}975$

$\frac{k}{\sigma} = 1{,}96 \quad k = 1{,}96 \cdot \sigma$

$[\mu - k, \mu + k] = [\mu - 1{,}96 \cdot \sigma, \mu + 1{,}96\sigma] = [120 - 1{,}96 \cdot 4, 120 + 1{,}96 \cdot 4] = [112{,}16, 127{,}84]$

Problem 3. (approx. 15%)
Wanting to optimise storage space a seller wants to model the number of orders on a specific product in December. In December the previous year the number of orders was 15.

1. Specify a random variable and its distribution, so that it describes that number of orders in December — explain your choice. $X \sim \text{Poisson}(15)$ (Tabel A.2)
2. What is the probability of 17 or more orders. $P(X \geq 17) = 1 - P(X \leq 16) = 1 - 0{,}6641 = 0{,}3359$
3. How large does stock need to be for the seller to have at least a 95% probability of fulfilling all orders? Assume that the seller cannot receive new stock during December.

$P(X \leq k) \geq 0.95$ $P(X \leq 22) = 0.9673$: Needs 22 items in stock.

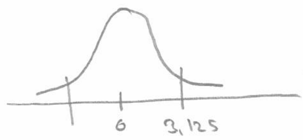
Problem 4. (approx. 30%)
The walls in a plastic bottle need to have a certain thickness to avoid that the bottle ~~does~~ breaks. An engineer in quality control takes a sample of 25 bottles and measures the wall thickness obtaining a sample average of $\bar{x} = 4.05mm$ and a sample standard deviation of $s = 0.08mm$. He further assumes that the observations are independent and normal distributed.

1) $(1-\alpha)100\%$ conf-int: $\bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} = 4.05 \pm 2.064 \cdot \frac{0.08}{\sqrt{25}} = [4.017; 4.083]$ (Table A.4)

1. Determine a 95% confidence interval for the mean of the wall thickness.
2. Determine a 95% confidence interval for the standard deviation of the wall thickness.
3. Test at the 5% significance level if the wall thickness is less than $4mm$.
4. Test at the 5% significance level if the standard deviation of the wall thickness equals 0.1

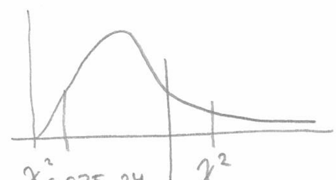
2) $\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} ; \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right] = \left[\frac{24 \cdot 0.08^2}{39{,}364} , \frac{24 \cdot 0.08^2}{12{,}401}\right] = [0{,}0039 ; 0{,}0124]$

3) $H_0: \mu \geq 4mm$
$H_1: \mu < 4$
$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{4.05 - 4}{0.08/\sqrt{25}} = 3{,}125$



$-t_{\alpha/2,n-1} = -1{,}71$
Do not reject H_0

4) $H_0: \sigma = 0.1$ $H_1: \sigma \neq 0.1$
$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{24 \cdot 0.08^2}{0.1^2} = 15.36$



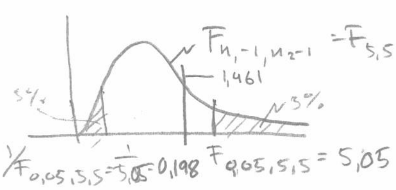
$\chi^2_{0{,}975,24} = 12.401$ $\chi^2_{0{,}025,24} = 39.364$
Do not reject H_0

Problem 5. (approx. 20%)
A cement factory wants to buy a new machine for filling bags with 50kg of cement. They have two machines to choose from. From each machine they take a sample of 6 bags and weigh each of them. The measured weight are given in the table below

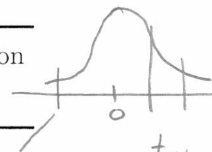
	Machine I	Machine II
	51.2	29.4
	49.0	50.7
	49.8	49.1
	51.7	48.7
	50.3	48.7
	51.4	50.1
$\bar{x}$	50.57	49.80
s^2	1.0987	0.7520

1. Test at the 10% significance level if the variance of the weights are equal for the two machines.
2. Test at the 10% significance level if the means of the weights are equal for the two machines.

1) $H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$
$f = \frac{s_1^2}{s_2^2} = \frac{1{,}0987}{0{,}7520} = 1{,}461$
Cannot reject H_0: Assume equal variance!



2) $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$
$Sp^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = 0{,}92535$
$t = \frac{\bar{x}_1 - \bar{x}_2}{Sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 1{,}3864$



$t_{\alpha/2,n_1+n_2-1} = t_{0{,}05,10} = 1{,}812$
Cannot reject H_0

Remember to add student number on all sheets and state how many sheets your solution consists of.

Written exam in Probability Theory and Statistics - K7

Lecturer: P. Svante Eriksen

Thursday 13th of january 2011, 9:00-13:00

In the assessment emphasis will be put on both correct methods as well as correct answers. Hence the method should be clearly stated.

Problem 1. (approx 20%)

The random variable X has a normal distribution with mean 4 and variance 25. 1. Mean: $4E(X)+6 = 4\cdot 4+6 = 22$, Var: $16\cdot\sigma_X^2 = 400$

1. Calculate the mean and variance of the variable $4X + 6$.
2. Calculate $P(0 \le X \le 4)$. $= P(X \geq 0) - P(X \geq 4) = 0.5 - 0.212 = 0.288$

The random variable Y has mean 5 and variance 10. The correlation coefficient of X and Y is -0.5. Mean: $4E(X) + 5E(Y) + 1 = 4\cdot 4 + 5\cdot 5 + 1 = 42$

3. Calculate the mean and variance of the variable $4X + 5Y + 1$. Var: $4^2\sigma_X^2 + 5^2\sigma_Y^2 + 2\cdot 4\cdot 5\,\sigma_{XY} =$
$16\cdot 25 + 25\cdot 10 + 40\times(-0.5\cdot\sqrt{25\cdot 10})$
≈ 333.8

Dist of X:

$x \backslash y$	0	1	2	SUM
0	0	$\frac{1}{27}$	$\frac{2}{27}$	$\frac{3}{27}$
1	$\frac{2}{27}$	$\frac{3}{27}$	$\frac{4}{27}$	$\frac{9}{27}$
2	$\frac{4}{27}$	$\frac{5}{27}$	$\frac{6}{27}$	$\frac{15}{27}$

Problem 2. (approx 20%)

The joint probability distribution of X and Y is given by

$$f(x, y) = \frac{2x + y}{27}, \quad x = 0, 1, 2; \; y = 0, 1, 2$$

1. Evaluate the marginal distribution of X. $P(Y=2 \mid X=1) = \frac{f(1,2)}{P(X=1)} = \frac{4}{9}$
2. Find $P(Y = 2|X = 1)$ and $P(Y = 2|X = 2)$. Are X and Y statistically independent? $P(Y=2 \mid X=2) = f(2,2)/P(X=2) = \frac{6}{15}$

No: $P(Y=2 \mid X=1) \neq P(Y=2 \mid X=2)$

3. Evaluate $E(X^2Y)$. $\sum_{x=0}^{2}\sum_{y=0}^{2} x^2 y\, f(x,y) = 1\cdot\frac{3}{27} + 2\cdot\frac{4}{27} + 4\cdot\frac{5}{27} + 8\cdot\frac{6}{27} = \frac{79}{27}$

Problem 3. (approx 10%)

In a certain city the need for money to buy drugs is stated as the reason for 60% of all thefts.

Consider the next 20 theft cases in the city and let X denote the number of cases resulting from the need for money to buy drugs.

1. Calculate the mean and variance of X. $\sim$ binomial $(n=20,\ p=0.6)$
2. Evaluate $P(4 \le X \le 12)$. $EX = n\cdot p = 12$ $\quad \sigma_X^2 = n\cdot p(1-p) = 4.8$

$= P(X \leq 12) - P(X \leq 3) \approx 0.548$

Problem 4. (approx 30%)
An engineer in quality control takes a sample of 30 bolts and measures their diameter, which yields a sample average of $\bar{x} = 10.023mm$ and a sample standard deviation $s = 0.009mm$. He assumes that the observations are a random sample from the normal distribution.

Walpole P. 279

$\nu = 29, \; t_{\alpha/2} = 2.045$

$\rightarrow [10.01964, 10.02635]$

1. Determine a 95% confidence interval for the mean of the bolt diameter.
2. Determine a 95% confidence interval for the standard deviation of the bolt diameter.
3. Test at the 5% significance level whether the bolts meet a requirement of a mean diameter equal to $10mm$.
4. Test at the 2.5% significance level whether the measurements meet a requirement of a standard deviation below or equal to $0.005mm$.

Not in confidence interval ⇒ reject.

Walpole p. 307:

$\nu = 29,$

$\chi^2_{\alpha/2} = 16.05$

$\chi^2_{1-\alpha/2} = 45.72 \rightarrow$

$[0.514, 1.464] \times 10^{-4}$

$\sqrt{\;} \rightarrow [0.717, 1.210] \times 10^{-2}$

Problem 5. (approx 20%)
Two methods for measuring the molar heat of fusion of water are being compared. Ten measurements made by method A have a sample mean $\bar{x}_A = 6.025$ kilojoules per mole and sample standard deviation of $s_A = 0.024KJ/mol$. Five measurements made by method B have a sample mean $\bar{x}_B = 6.001KJ/mol$ and sample standard deviation of $s_B = 0.012KJ/mol$.

1. Test at the 5% significance level whether the two methods have the same standard deviation.
2. Test at the 5% significance level whether the mean measurements differ between the two methods.

Walpole p 368

$F = \left(\frac{24}{12}\right)^2 = 4$

$\nu_1 = 9 \quad \nu_2 = 4$

$f_{\alpha/2}(9,4) = 8.9 > 4$

Do not reject

Remember to add student number on all sheets and state how many sheets your solution consists of

---

Walpole p. 346: $s_p^2 = 4.43 \times 10^{-4}$

$t = 2.082 < t_{\alpha/2, 13} = 2.16$

Do not reject.