MIDTERM 2 (MATH 4653, FALL 2008)

Your Name: ________________________________

Date: ______________________

The rules:
You are allowed to use only this paper and only the most basic calculator.
No books, no notebooks, no web access. You MUST write your name.
You MUST simplify and box all answers. Except for the last problem,
you MUST write out all intermediate calculations and logical reasoning.
You can use formulas and the table on the last two pages.
You have exactly 100 minutes.

Points:

1	
2	
3	
4	
5	
6	

____

Total: **(out of 100)**

Problem 1. (15 points)
Cookies produced by a factory in Brussels are being tested, to decide whether their weight confirms with the standards. For a batch of 100 cookies the results are as follows: 9.7 - 4 times, 9.8 - 14 times, 9.9 - 22 times, 10 - 34 times, 10.1 – 16 times, 10.2 - 8 times, 10.3 - 2 times (all in grams).

a) Find the 90% and 95%-confidence intervals for the cookie weights.

b) The factory claims that the average cookie weights 10 with s.d. 0.1. Should we believe the factory or are they shortchanging the customers? (use the $> 95\%$ confidence)

c) The consumer advocate group reported that the average weight of a cookie is really 9.9 with s.d. 0.1, so the factory is misleading at best (off the record, they claim that the factory CEO is a lying liar). Should we believe the consumer advocates?

Problem 2. (15 points)
Casino has the following game: 4 dies are thrown, the sum is calculated and if the sum is ≥ 22, a player wins a prize. Exactly 10,000 games were played in one year.

a) Use the CLT to estimate the probability of that there are more than 100 wins.

b) Use the CLT to estimate the probability of that there are less than 120 wins.

c) Use the CLT to estimate the probability of that the number wins is between 110 and 150.

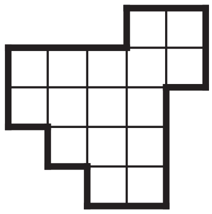
Problem 3. (15 points)
A fair die is rolled 900 times. Denote by X the sum of outcomes.

$a)$ Use Markov inequality to estimate the probability that $X > 6000$.

$b)$ Use Chebyshev inequality to estimate the probability that $X > 3300$.

$c)$ Use CLT to estimate the probability that $X > 3300$.

Problem 4. (13 points)
Suppose (X, Y) is a random integer point in the following figure. Compute $\rho(X, Y)$.



Problem 5. (15 points)
A townhall has a weekly bingo game which the same 100 people attend. The game was on 50 times a year, for 10 years.

a) Is it unusual if every single person won at least once? (i.e. decide whether the probability of that is $> 5\%$).

b) Is it unusual if the town mayor (who has been a mayor for the last 20 years and was just reelected again) won at least 10 times?

c) Is it unusual if at least one person in town won at least 20 times?

Problem 6. (27 points, 3 points each). **True or False?**
Give answers only, next to the questions. No reasoning/calculations will be taken into account.

a) When a null hypothesis H_0 is tested and the confidence with which it must be rejected increased from 95% to 98%, the type I error is increased.

b) In the $100(1-\alpha)$%-confidence interval, when α is increased the confidence interval also increases.

c) The square root law gives a formula for the mean of the average of n i.i.d. random variables $X_1, \ldots, X_n$.

d) The CLT gives only the approximate, not exact values for the probabilities.

e) In a large Midwestern city in 2007 there were 900 car fatalities compared to average of 850. You are told that the number fatalities behaves according to the Poisson distribution and that the CLT is to be used. Should you use mean 900 and s.d. 900?

f) $Var(X-Y) = Var(X) - Var(Y)$ for all independent X, Y.

g) If $\rho(X, Y) = 0$, then X and Y are independent.

h) Two polls are made on the popularity of the *Circle* toothpaste. One called random 1,000 people and found 35% popularity. Another called 4,000 people and found 40% popularity. A company concluded that its popularity is roughly 39%. True or false: you should agree.

i) A pharmaceutical company reported 90%-confidence intervals for the quality of its new miracle pill, while kept all the data secret. It only disclosed that the number of clinical study participants is large, since this is an FDA requirement. Then a statistician working for a competing company can obtain the 98%-confidence interval for this pill, so to compare with their own pill in development.

Bernoulli distribution Ber(p)
$P(X = 1) = p$, $P(X = 0) = 1 - p$, $E[X] = p$, $\mathrm{Var}(X) = p(1 - p)$.

Binomial distribution Bin(n, p)
$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $E[X] = np$, $\mathrm{Var}(X) = np(1 - p)$.

Geometric distribution Geo(p)
$P(X = k) = p(1 - p)^{k-1}$, $E[X] = 1/p$, $\mathrm{Var}(X) = (1 - p)/p^2$.

Poisson distribution Poi(λ)
$P(X = k) = e^{-\lambda} \lambda^k / k!$, $E[X] = \lambda$, $\mathrm{Var}(X) = \lambda$.