This is just a set of problems and not an assignment. You might find some problems interesting some having appeared part of past assignments, some coming in the future assignments etc... If you solve a particularly hard problem, you can come to my office to proudly claim extra credits.

i. This ~~assignment~~ set has in total **46** problems split into two sections and further into five sub-sections.

ii. None of the problems is particularly hard but all problems would help in better understanding of the material dealt with in-class.

iii. Effort has been made to minimize repetitive ideas among the problems and would appreciate students pointing out the remaining such occurrences.

iv. The computing problems should be done in a systematic way on a worksheet application, preferably using Microsoft Excel. The worksheet/workbook need to be turned in only at the start of the finals week.

v. ~~The problems other than the computing problems are due on the **25th** of September 2000.~~

vi. Hope you have fun. Good Luck!

Prerequisites

Probability I

Problem 1 (Independence and Covariance). In the following $B(1, 0.5)$ is a symmetric Bernoulli random variable.

1. Let $X = 2(B(1, 0.5) - 0.5)$ and $Y = X^2$.

   a. Show that $\text{Cov}(X, Y) = 0$.

   b. Show that the two random variables are independent.

2. Let X be such that

$$\text{P}\{X = x\} = \begin{cases} 0.25 & x = -2, -1, 1 \text{ and } 2 \\ 0 & \text{otherwise} \end{cases}$$

   and $Y = X^2$.

   a. Show that $\text{Cov}(X, Y) = 0$.

   b. Show that the two random variables are dependent.

Problem 2 (Simulation). This problem explains why computers need to only be able to generate uniform random variables. Let F be a distribution function.

a. Define $\text{F}^{-1}(u) = \inf\{z|\text{F}(z) \geqslant u\}$. Prove that $\{u|\text{F}^{-1}(u) \leqslant x\} = (-\infty, \text{F}(x)]$, $\forall u \in (0, 1)$
   Hint :

$$\begin{aligned} u \leqslant \text{F}(x) &\Rightarrow x \in \{z|\text{F}(z) \geqslant u\} \\ &\Rightarrow x \geqslant \inf\{z|\text{F}(z) \geqslant u\} \\ &\iff x \geqslant \text{F}^{-1}(u),\ \forall u \in (0, 1) \end{aligned}$$

b. Let U be a $\mathrm{U}(0,1)$ distributed random variable. Show that $\mathrm{F}^{-1}(U)$ has F as it's distribution function.

c. Why did we have to define the *inverse* of F?

Problem 3 (Negative Binomial). Consider a coin which has probability p of landing *heads-up*.

a. Show that if X denotes the number of tosses to the first head then X has a geometric distribution.

b. Find the distribution of the number of tosses to the second head.

c. Generalize above for the n*th* head. This former distribution is a negative binomial distribution.

d. Conclude from above that if X and Y are independent geometric random variables with the same expectation, then $X+Y$ has a negative binomial distribution.

Problem 4 (Existence of Moments). Let X be a random variable taking values in $\{1, 2, \ldots\}$ such that

$$\mathrm{P}\{X = n\} \propto \begin{cases} \frac{1}{n^2 \log(n)(\log(\log(n)))^2} & n = 5, 6, \ldots \\ 0 & \text{otherwise} \end{cases}$$

a. Show that $\mathbb{E}(X) < \infty$

b. Show that $\mathbb{E}\big(X^{1+\epsilon}\big) = \infty, \quad \forall \epsilon > 0$

Problem 5 (On Variance). In the following assume the existence of all required moments.

i. It is easy to see that

$$\mathrm{Var}\,(X) = \min_{a \in \mathbb{R}} \mathbb{E}(X-a)^2$$

  a. Prove using Calculus
  b. Prove without using Calculus

ii. Let X and Y be two random variables.

  a. Show that $X - \mathbb{E}(X|Y)$ is uncorrelated with $\mathbb{E}(X|Y)$
  b. Hence or otherwise show that

$$\mathrm{Var}\,(X) = \mathbb{E}(X - \mathbb{E}(X|Y))^2 + \mathbb{E}(\mathbb{E}(X|Y) - \mathbb{E}(X))^2$$
$$\mathrm{Var}\,(X) = \mathbb{E}(\mathrm{Var}\,(X|Y)) + \mathrm{Var}\,(\mathbb{E}(X|Y))$$

  c. Can you see it as a **Pythagoras Theorem**.

Problem 6 (On Mean and Median). The following give an alternate definition of the mean and median.

i. Show that the mean is the value of a which minimizes $\mathbb{E}(X-a)^2$

ii. Show that the set of values where $\mathbb{E}(|X-a|)$ is minimized is the set of medians of the distribution of X. Can you specify conditions when the median is unique. Note that a median is any real, m, which satisfies the condition

$$\Pr(X \geqslant m) \wedge \Pr(X \leqslant m) \geqslant 0.5$$

iii. Show that the MEAN DEVIATION, $\mathbb{E}(|X-\mathbb{E}(X)|)$, satisfies

$$\mathbb{E}(|X-\mathbb{E}(X)|) \leqslant \sqrt{\operatorname{Var}(X)}$$

iv. Using the above results, show that

$$|\mathbb{E}(X) - \text{median}\{X\}| \leqslant \sqrt{\operatorname{Var}(X)}$$

Financial Mathematics I

Problem 1. Consider the amount function $A(\cdot)$ given by

$$A(t) = \frac{1}{\log_5(2t+10) - \log_5(2t+8)}, \quad t \geqslant 0$$

a. Find the corresponding accumulation function $a(\cdot)$.

b. Find the corresponding value of the annuity $a_{\overline{n}|}, \quad n \geqslant 1$.

c. Find I_n, the amount of interest paid during the n*th* period.

Problem 2. *Stoodley's formula* for the force of interest is

$$\delta(t) = p + \frac{s}{1+re^{st}}, \quad t \geqslant 0$$

a. Show that the corresponding accumulation function $a(t)$ is such that

$$\frac{1}{a(t)} = \frac{1}{1+r} \cdot \nu_1^t + \frac{r}{1+r} \cdot \nu_2^t, \quad t \geqslant 0$$

where $\nu_1 = e^{-(p+s)}$ and $\nu_2 = e^{-p}$

b. Find the corresponding value of the annuity $a_{\overline{n}|}, \quad n \geqslant 1$.

c. Find I_n, the amount of interest in the n*th* period.

Problem 3. Prove that

a.

$$\sum_{m=1}^{\infty} (-1)^{m-1} i^m \left[\frac{1}{d^{(m)}} - \frac{1}{i^{(m)}} \right] = \delta$$

b.

$$\frac{d^n}{d\nu^n}(\delta \nu^{n-1}) = -(1+i)(n-1)!$$

c.

$$\frac{d}{dt}(\delta(t)) = \frac{A''(t)}{A(t)} - \delta(t)^2$$

Problem 4. Show

a.

$$\frac{\ddot{s}_{\overline{2n}|}}{\ddot{s}_{\overline{n}|}} + \frac{\ddot{s}_{\overline{n}|}}{\ddot{s}_{\overline{2n}|}} - \frac{\ddot{s}_{\overline{3n}|}}{\ddot{s}_{\overline{2n}|}} = 1$$

b.

$$\sum_{t=10}^{15} (\ddot{s}_{\overline{t}|} - s_{\overline{t}|}) = s_{\overline{16}|} - s_{\overline{10}|} - 6$$

Problem 5. All the following could be *visualized*.

a. If $\ddot{a}_{\overline{p}|} = x$ and $s_{\overline{q}|} = y$, show that

$$a_{\overline{p+q}|} = \frac{\nu x + y}{1 + iy}$$

b. Find x, y and z if

$$\frac{a_{\overline{9}|}}{a_{\overline{12}|}} = \frac{a_{\overline{4}|} + s_{\overline{x}|}}{a_{\overline{y}|} + s_{\overline{z}|}}$$

c. Simplify $a_{\overline{12}|} \cdot (1 + \nu^{12} + \nu^{24} + \nu^{36})$

<u>Survival Distributions and Life Tables</u>

<u>Textbook Problems</u>

The following refer to problems from *Actuarial Mathematics*, **2nd** Edition by Bowers et al.

#	1	2	3	4	5	6	7	8	9	10
Problem # in Text	16	17	19	21	22	24	25	35	39	40

<u>Other Theoretical Problems</u>

Problem 1 (Laws with Constant Density/Mortality Rate). For the below given distributions of X find the following for $X, T(x)$ and $K(x)$.

i. Mean and Variance

ii. Density function, Distribution function and Survival function

iii. Mortality Rate

List of Distributions:

a. Let $\mathrm{U}(a,b)$ represent the uniform distribution on the interval (a,b) where $-\infty < a < b < \infty$.

b. Let $\mathrm{E}(\lambda)$ represent the exponential distribution with $\lambda > 0$

Problem 2 (Analytical Mortality Laws). Show that the below given survival functions satisfy the properties of a survival function and moreover that the given forces of mortality are the corresponding ones. Also, find the set of values of x that maximizes the value of $S(x)\mu(x)$.

Originator	$S(x)$	$\mu(x)$
De Moivre	$1-\frac{x}{\omega}, \quad 0 \leqslant x \leqslant \omega$	$\frac{1}{(\omega-x)}$
Gompertz	$e^{[-m(c^x-1)]}, \quad x \geqslant 0, c > 1, B > 0$	Bc^x
Makeham	$e^{[-Ax-m(c^x-1)]}, \quad x \geqslant 0, c > 1, A \geqslant -B, B > 0$	$A + Bc^x$
Weibull	$e^{[-ux^{n+1}]}, \quad x \geqslant 0, n > 0, k > 0$	kx^n
Pareto	$\frac{b^a}{(b+x)^a}, \quad x \geqslant 0, a > 0, b > 0$	$\frac{a}{b+x}$

where $m = \frac{B}{\ln(c)}$ and $u = \frac{k}{n+1}$.

Problem 3 (Estimation of $\mathrm{e}_{0:\overline{n}|}$). Let the survival function of **age-at-death** follow $\mathrm{U}(0,80)$ and l_0 be equal to 10. Below, ${}_{\mathrm{n}}\mathcal{D}_{\mathrm{x}}$ would denote the number of death in the cohort between ages x and x+n.

a. Find the distribution of $({}_1\mathcal{D}_0\ {}_1\mathcal{D}_1,\ {}_1\mathcal{D}_2,\ {}_{77}\mathcal{D}_3)$. Is it a *standard* distribution?

b. Find the mean and variance of each of the above random variables?

c. What are the correlations and covariances between any pair of them?

d. Show that it is *sensible* to estimate $\mathrm{e}_{0:\overline{3}|}$ by $\widehat{\mathrm{e}_{0:\overline{3}|}} = \sum_{i=1}^{2} i\ {}_1\mathcal{D}_i + 3\ {}_{77}\mathcal{D}_3$ by find the expectation of the latter. Moreover, find the variance of this estimate.

Problem 4 (Graphical Properties). Assume in the following, any differentiability or continuity of functions that you may require.

a. Given that l_{x+t} is strictly decreasing as a function of t in the interval $[0,1]$ show the following.

  i. l_{x+t} is concave implies $\mathrm{q}_x \geqslant \mu(x)$
  ii. l_{x+t} is convex implies $\mathrm{q}_x \leqslant \mu(x)$

b. Prove that

$$\frac{\mathrm{d}}{\mathrm{dx}} l_x\mu(x) \gtreqqless 0 \quad \text{iff} \quad \frac{\mathrm{d}}{\mathrm{dx}} \mu(x) \gtreqqless \mu(x)^2$$

In the following twelve problems we shall be using the following notations.

Notation 1. Consider the one-parameter class of distributions, $\left\{f^{\mathrm{H}}_\alpha | \alpha > 0 \text{ but } \alpha \neq 1\right\}$ where

$$f^{\mathrm{H}}_\alpha(x) = \begin{cases} 0 & x < 0 \\ \frac{\alpha}{(\alpha+(1-\alpha)x)^2} & 0 \leqslant x \leqslant 1 \\ 0 & x > 1 \end{cases}$$

All probabilistic functions/quantities with superscript H and a positive parameter would relate to this class of distributions.

Notation 2. Consider the one-parameter class of distributions, $\left\{f_\lambda^{\mathrm{E}} | \lambda > 0\right\}$ where

$$f_\lambda^{\mathrm{E}}(x) = \begin{cases} 0 & x < 0 \\ \frac{\lambda e^{-\lambda x}}{(1-e^{-\lambda})} & 0 \leqslant x \leqslant 1 \\ 0 & x > 1 \end{cases}$$

All probabilistic functions/quantities with superscript E and a positive parameter would relate to this class of distributions.

Notation 3. All probabilistic functions/quantities with superscript U would relate to the uniform distribution on the interval $[0, 1]$.

Problem 5. For the above two families find the following quantities.

i. Mean and Variance

ii. Distribution and Survival functions

iii. Mortality rate

Problem 6 (Convergence). For both the families find the limits of the following quantities as the α approaches one and λ approaches zero, respectively.

1. Distribution and Survival function
2. Density function
3. Mortality Rate
4. Mean and Variance and all moments

Hence, what would be the natural definition for f_1^{H} and f_0^{E}?

Problem 7 (Sources of Origin). Show the following.

1. The above family can be derived from the exponential distribution by considering the conditional distribution given $X < 1$
2. The above family can be derived from the exponential distribution by considering the decimal part of the random variable. What would be the distribution of the integer part?

Problem 8 (Stochastic Ordering - Family I). The following give an interesting perspective of the first family.

1. Assume $0 < \alpha < 1$. Show that

$$F_\alpha^{\mathrm{H}}(t) - F^{\mathrm{U}}(t) \text{ is } \begin{cases} > 0 & t \in (0, 1) \\ = 0 & \text{otherwise} \end{cases}$$

2. Assume $\alpha > 1$. Show that

$$F_\alpha^{\mathrm{H}}(t) - F^{\mathrm{U}}(t) \text{ is } \begin{cases} < 0 & t \in (0,1) \\ = 0 & \text{otherwise} \end{cases}$$

3. What about $\alpha = 1$?

Problem 9 (Stochastic Ordering - Family II). It is a standard result from calculus that

$$\boxed{e^x - (1+x) \text{ is } \begin{cases} > 0 & |x| > 0 \\ = 0 & x = 0 \end{cases}}$$

Use the above to show that, for all $\lambda > 0$

$$F_\alpha^{\mathrm{E}}(t) - F^{\mathrm{U}}(t) \text{ is } \begin{cases} > 0 & t \in (0,1) \\ = 0 & \text{otherwise} \end{cases}$$

What about the case $\lambda = 0$?

In the following problems we will assume the existence of a Life/Mortality table, which would give us the values of q_x for all non-negative integer values of x. Hence the survival distribution function of X is determined at the integer values. We shall use examine in the following the different assumptions that is used to determine the values of the survival function at non-integer values.

Notation 4. Note that we could split the age-at-death random variable X as

$$X = \lfloor X \rfloor + (X - \lfloor X \rfloor)$$

where

$$\lfloor X \rfloor = sup\,\{k | k \leqslant X, k \in \mathcal{I}\}$$

We shall use the following notations in the following.

$$K = \lfloor X \rfloor$$
$$F = X - \lfloor X \rfloor$$

Problem 10 (Uniform Assumption). Show that the following are equivalent.

i. Linear interpolation of the survival function.

ii. K and F are independent and $F \sim \mathrm{U}(0,1)$

Problem 11 (Constant Force Assumption). Show that the following are equivalent.

i. *Geometric interpolation* of the survival function.

ii. The conditional density of F given $K = k$ is $f^{\mathrm{E}}_{-\log \mathrm{p}_k}$

Problem 12 (Hyperbolic Assumption). Show that the following are equivalent.

i. *Harmonic interpolation* of the survival function.

ii. The conditional density of F given $K = k$ is $f^{\mathrm{H}}_{\mathrm{p}_k}$

Problem 13 (Independence of K and F). Show the following:

i. K is independent of F if and only if $\frac{{}_t\mathrm{q}_k}{\mathrm{q}_k}$ as a function of k and t does not depend on k.

ii. Hence or otherwise conclude that under the uniform assumption for fractional ages we have the above independence but not under the other two assumptions.

Problem 14 (Stochastic Ordering - Revisited). Above, we saw that the uniform distribution stochastically dominates both the constant force and hyperbolic assumptions. Note that we used curve tracing to show this ordering. Now using the above characterizations of the uniform, constant force and hyperbolic assumption, show that the uniform dominates the constant force which in turn dominates the hyperbolic assumption.

Problem 15 ($a(x)$ and Constant Force). Let $a(k)$ denote as usual, $\mathbb{E}(T(k)|T(k) < 1)$. Show the following under the constant force assumption;

i. If $\mu = -\log \mathrm{p}_k$, then

$$a(k) = \frac{1 - e^{-\mu} - \mu e^{-\mu}}{\mu(1 - e^{-\mu})}$$

ii. It must be clear from above that $\lim_{\mathrm{q}_x \to 0} a(k) = 0.5$. Moreover,

$$\lim_{\mathrm{q}_x \to 0} \frac{a(k) - 0.5}{\mathrm{q}_x} = -\frac{1}{12}$$

Problem 16 ($a(x)$ and Hyperbolic). Let $a(k)$ denote as usual, $\mathbb{E}(T(k)|T(k) < 1)$. Show the following under the constant force assumption;

i.

$$a(k) = -\frac{\mathrm{p}_k}{{\mathrm{q}_x}^2}(\mathrm{q}_x + \log(\mathrm{p}_x))$$

ii. It must be clear from above that $\lim_{\mathrm{q}_x \to 0} a(k) = 0.5$. Moreover,

$$\lim_{\mathrm{q}_x \to 0} \frac{a(k) - 0.5}{\mathrm{q}_x} = -\frac{1}{6}$$

iii. Can you rationalize the above limits with what we saw in the stochastic ordering problem above?

Problem 17 (A Useful Table). In the following table, $0 \leq t + y \leq 1$ and x is an integer.

In terms of	${}_t\mathrm{q}_x =$	$\mu(x+t) =$	${}_y\mathrm{q}_{x+t} =$	${}_t\mathrm{p}_x =$	${}_t\mathrm{p}_x\mu(x+t) =$
${}_t\mathrm{q}_x$	.	$-\frac{\mathrm{d}}{\mathrm{dt}}\ln(1-{}_t\mathrm{q}_x)$	$\frac{{}_{y+t}\mathrm{q}_x-{}_t\mathrm{q}_x}{1-{}_t\mathrm{q}_x}$	$1-{}_t\mathrm{q}_x$	$\frac{\mathrm{d}}{\mathrm{dt}}\,{}_t\mathrm{q}_x$
$\mu(x+t)$	$1-e^{-\int_0^t \mu(x+z)\mathrm{d}z}$	.	$1-e^{-\int_t^{t+y} \mu(x+z)\mathrm{d}z}$	$e^{-\int_0^t \mu(x+z)\mathrm{d}z}$	$\mu(x+t)e^{-\int_0^t \mu(x+z)\mathrm{d}z}$
${}_y\mathrm{q}_{x+t}$	Changing t to 0 and y to t, we get ${}_t\mathrm{q}_x$. Then we proceed as with ${}_t\mathrm{q}_x$.				
${}_{1-t}\mathrm{q}_{x+t}$	$\frac{\mathrm{q}_x-{}_{1-t}\mathrm{q}_{x+t}}{1-{}_{1-t}\mathrm{q}_{x+t}}$	$\frac{\mathrm{d}}{\mathrm{dt}}\ln(1-{}_{1-t}\mathrm{q}_{x+t})$	Through ${}_t\mathrm{q}_x$	$\frac{1-\mathrm{q}_x}{1-{}_{1-t}\mathrm{q}_{x+t}}$	Use ${}_t\mathrm{p}_x\mu(x+t)$
${}_t\mathrm{p}_x$	$1-{}_t\mathrm{p}_x$	$-\frac{\mathrm{d}}{\mathrm{dt}}\ln({}_t\mathrm{p}_x)$	$1-\frac{{}_{y+t}\mathrm{p}_x}{{}_t\mathrm{p}_x}$	.	$-\frac{\mathrm{d}}{\mathrm{dt}}\,{}_t\mathrm{p}_x$
${}_t\mathrm{p}_x\mu(x+t)$	$\int_0^t {}_z\mathrm{p}_x\mu(x+z)\mathrm{d}z$	Use $\frac{{}_t\mathrm{p}_x\mu(x+t)}{{}_t\mathrm{p}_x}$	$\frac{\int_t^{t+y} {}_z\mathrm{p}_x\mu(x+z)\mathrm{d}z}{\int_t^{\infty} {}_z\mathrm{p}_x\mu(x+z)\mathrm{d}z}$	$\int_t^{\infty} {}_z\mathrm{p}_x\mu(x+z)\mathrm{d}z$	.

i. Derive the above table.

ii. For the three fractional age assumptions discussed above, find the left and right limits of the mortality rate at integral points using the above table.

iii. For the hyperbolic assumption, looking at the formula for ${}_t\mathrm{p}_x$, justify its name.

Problem 18 (Proportional Hazard Model). Let us denote by $\mathbf{x}$, the vector of information about the life at the time of selection and μ denote the standard force of mortality. Then in the proportional hazard model, there exists a function Ψ such that

$$\mu_{\mathbf{x}}(t) = \psi(\mathbf{x})\mu(t), \quad \forall t \geqslant 0$$

It is clear that if $\mathbf{x}_0$ represents the *standard* information, then $\Psi(\mathbf{x}_0) = 1$. Show that

i. ${}_t\mathrm{p}_{[\mathbf{x}]} = ({}_t\mathrm{p}_{[\mathbf{x}_0]})^{\Psi(\mathbf{x})}$.

ii. Density of $T(\mathbf{x})$ is given by

$$-\Psi(\mathbf{x})({}_t\mathrm{p}_{[\mathbf{x}_0]})^{\Psi(\mathbf{x})-1}\frac{\mathrm{d}_t\mathrm{p}_{[\mathbf{x}_0]}}{\mathrm{d}t}.$$

iii. If $\Psi^\circ = 2\Psi$, then prove that

$${}_t\mathrm{q}_{[x]}{}^{\Psi^\circ} \leqslant 2\,{}_t\mathrm{q}_{[x]},$$

where ${}_t\mathrm{q}_{[x]}{}^{\Psi^\circ}$ is ${}_t\mathrm{q}_{[x]}$ calculated using Ψ°.

iv. For a general model, does a similar relationship as above exist between the ${}_t\mathrm{q}_x$'s after the mortality rate is doubled?

Problem 19 (Central Death Rate). Find the central death rate, m_x, in terms of q_x under each of the following assumptions.

1. Uniform Assumption
2. Constant Force
3. Hyperbolic Assumption

Problem 20. If instead of $\mathrm{q}_{100} = 1$ we choose $\mathrm{q}_x = 0.25, \quad \forall x \geqslant 100$, find the **increase** in the expected number of years lived by a person with age 100, using the uniform fractional assumption.

Computing Problems

Problem 1 (Life Tables). Using a worksheet application, do the following.

i. Generate life tables similar to the handout on **AMA 91** for the **CNSF 2000** and **México 2000** mortality tables.

ii. For the CIA basic mortality tables found on the course web site, find the corresponding life tables assuming $l_{[0]} = 100000$.

iii. Plot the survival functions for the six tables using the three fractional age assumptions on the ultimate rates.

iv. Using the backward recurrence relationship for $e_{x:\overline{n}|}$, derive the values of $e_{x:\overline{10}|}$ for the **CNSF 2000** tables.

v. Using the backward recurrence relationship for $\mathring{e}_x$ and the trapezoidal rule, derive the values of $\mathring{e}_x$ for the **México 2000** tables.

Problem 2 (Index of Selection). The quantity

$$I(x,k) \stackrel{\text{def}}{=} 1 - \frac{{}_k q_{[x]}}{q_{x+k}}$$

For the CIA basic mortality tables found on the course web site, graph the function I using a 3-dimensional plot. What does it signify?

Problem 3. Using uniform, constant force and hyperbolic assumption for fractional age find the expected number of years that a student who joins I.T.A.M. at aged 17 would spend in I.T.A.M. if all the students graduate in exactly four years. Use the **México 2000** female table.

Problem 4. In ITAM, 1200 students joined in the first year of different programs. Assuming that they take 4 years to graduate, find the number of expected deaths among these students if the age distribution is as given below. Use the **CNSF 2000** group table.

x	$Number of Students$
17	10%
18	60%
19	20%
20	10%

Problem 5 (Application of CLT). Consider a random survivorship group consisting of four sub-groups;

1. 500 newborns
2. 500 joining at age 25
3. 500 joining at age 50
4. 500 joining at age 75

Using the **México 2000** male table and further assuming that the lives are independent, arrive at an approximate value c satisfying

$$\Pr(Y_1 + Y_2 + Y_3 + Y_4 > c) = 0.05$$

where Y_i is the number of survivors from sub-group i till age 90.