Examiner: Xiangfeng Yang (013-285788). **Things allowed**: a calculator, a self-written A4 paper (two sides).
Scores rating (Betygsgränser): 8-11 points giving rate 3; 11.5-14.5 points giving rate 4; 15-18 points giving rate 5.
Notation: 'A random variable X is distributed as...' is written as '$X \in$ ... or $X \sim$ ... '

1 (3 points)

Let $X \sim U(0,1)$ and $Y \sim U(0,2)$ be two independent uniform random variables. Find the density function of $U = X + Y$. (Hint: You can use either convolution formula or transformation theorem. Be really really careful with the bounds of each variable!!! It might help to draw a graph for the bounds)

Solution. Let $U = X + Y$ and $V = X$. Then it is important to notice that $0 < V < 1$ and $0 < V < U < 2 + V < 3$ (these can be seen by noticing that $X = V \in (0,1)$ and $Y = U - V \in (0,2)$). Furthermore,

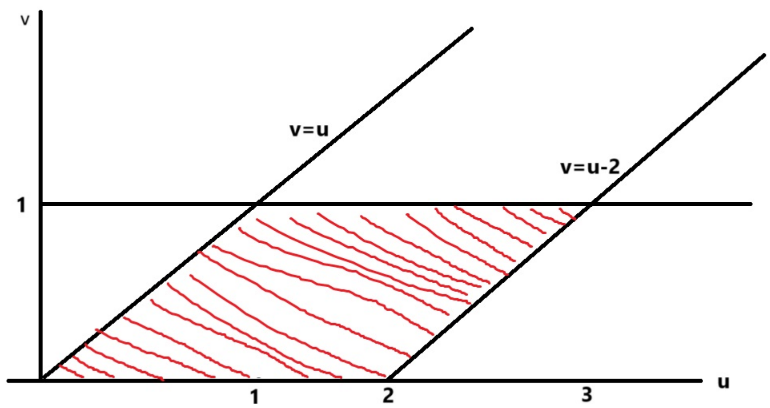
$$X = V, \quad Y = U - V, \qquad J = |\frac{\partial(x \;\; y)}{\partial(u \;\; v)}| = -1.$$

Therefore the joint probability density function of $(U,V)'$ is

$$f_{U,V}(u,v) = f(x^{-1}(u,v), y^{-1}(u,v))|J| = f_X(v)f_Y(u-v)|J|$$
$$= 1 \cdot \frac{1}{2} \cdot 1 = \frac{1}{2}, \text{ for } 0 < v < 1 \text{ and } 0 < v < u < 2 + v < 3.$$

In order to obtain the density function of U, we need to integrate with respect to v. Theretofore, it is important to know the bounds of v in terms of u. If we rewrite the non-trivial domain: $0 < v < 1$ and $0 < v < u < 2 + v < 3$ as follows

$$0 < u < 3, \qquad \max\{0, u-2\} < v < \min\{u, 1\},$$



Then it is clear that

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)dv = \int_{\max\{0,u-2\}}^{\min\{u,1\}} \frac{1}{2}dv = \frac{1}{2}\left(\min\{u,1\} - \max\{0,u-2\}\right), \quad \text{for } 0 < u < 3$$
$$= \begin{cases} \frac{1}{2}u, & \text{if } 0 < u < 1, \\ \frac{1}{2}, & \text{if } 1 \le u < 2, \\ \frac{1}{2}(3-u), & \text{if } 2 \le u < 3. \end{cases}$$

□

2 (3 points)

Let us throw a fair die twice independently. Set U = the outcome of the first throw and V = the outcome of the second throw. Define

$$X = U \qquad \text{and} \qquad Y = U + V.$$

Find the conditional expectation $E(Y|X = x)$.

Solution. The conditional probability mass function is

$$p_{Y|X=x}(y) = P(Y = y|X = x) = \frac{p_{X,Y}(x,y)}{p_X(x)} = \frac{\frac{1}{6}\cdot\frac{1}{6}}{\frac{1}{6}} = \frac{1}{6}$$

for any $x = 1,2,3,4,5,6$ and $y = x+1, x+2, x+3, x+4, x+5, x+6$. Therefore the conditional expectation is

$$E(Y|X = x) = \sum_y p_{Y|X=x}(y)\cdot y = \sum_{k=1}^{6} p_{Y|X=x}(x+k)\cdot(x+k) = \sum_{k=1}^{6}\frac{1}{6}\cdot(x+k) = x+3.5$$

for $x = 1,2,3,4,5,6$.

□

3 (3 points)

Consider the following situation: Hanna has a coin with $P(\text{head}) = p_1$ and Livia has a coin with $P(\text{head}) = p_2$. Hanna tosses her coin m times. Each time Hanna obtains "head", Livia tosses her coin (otherwise not). Let X be the total number of heads obtained by Livia. Then X can be modeled as follows:

$$X|N = n \sim Bin(n, p_2), \quad \text{with } N \sim Bin(m, p_1), \quad 0 < p_1, p_2 < 1,$$

where N denotes the total number of heads obtained by Hanna. Find the probability generating function (PGF) of X. Do you recognize the distribution of X?
(Hint: probability generating function of a Binomial random variable is $g_{Bin(n,p)}(t) = (q+pt)^n$ with $q = 1-p$)

Solution. The PGF of X can be computed as

$$\begin{aligned}
g_X(t) &= E(t^X) = E(E(t^X|N)) = E((q_2+p_2t)^N) \\
&= [q_1 + p_1(q_2+p_2t)]^m, \qquad (\text{where } q_1 = 1-p_1 \text{ and } q_2 = 1-p_2) \\
&= [(q_1+p_1q_2) + p_1p_2t]^m \\
&= [(1-p_1p_2) + p_1p_2t]^m \\
&= g_{Bin(m,\ p_1p_2)}(t).
\end{aligned}$$

Therefore $X \sim Bin(m,\ p_1p_2)$.

□

4 (3 points)

Let $X_1, X_2, \ldots, X_n$ be i.i.d. $Exp(1)$ random variables, and $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ be the order statistic. Define

$$Y_1 = X_{(1)}, \quad Y_k = X_{(k)} - X_{(k-1)}, \text{ for } k = 2,3,\ldots,n.$$

(4.1) (1p) Find the joint density function $f_{X_{(1)},X_{(2)},\ldots,X_{(n)}}(x_1,x_2,\ldots,x_n)$ of $(X_{(1)},X_{(2)},\ldots,X_{(n)})$.
(4.2) (1p) Find the joint density function $f_{Y_1,Y_2,\ldots,Y_n}(y_1,y_2,\ldots,y_n)$ of $(Y_1,Y_2,\ldots,Y_n)$.
(4.3) (1p) find the density function $f_{Y_n}(y_n)$ of Y_n.

Solution. (4.1) It is directly from Theorem 4.3.1 (book) that the joint density is

$$f_{X_{(1)},X_{(2)},\ldots,X_{(n)}}(x_1,x_2,\ldots,x_n) = n!f(x_1)f(x_2)\ldots f(x_n) = n!e^{-(x_1+x_2+\ldots+x_n)}, \qquad \text{for } 0 < x_1 < x_2 < \ldots < x_n.$$

(4.2) The following transform (with $Y_i > 0$)

$$Y_1 = X_{(1)}, \quad Y_k = X_{(k)} - X_{(k-1)}, \text{ for } k = 2,3,\ldots,n,$$

gives that

$$X_{(1)} = Y_1, \quad X_{(2)} = Y_1+Y_2, \qquad \ldots \qquad X_{(n)} = Y_1+Y_2+\ldots+Y_n.$$

Therefore the Jacobian is $J = 1$. This implies that the joint density is: for $y_i > 0$,

$$\begin{aligned} f_{Y_1,Y_2,\ldots,Y_n}(y_1,y_2,\ldots,y_n) &= f_{X_{(1)},X_{(2)},\ldots,X_{(n)}}(y_1, y_1+y_2, \ldots, y_1+y_2+\ldots+y_n)\cdot|J| \\ &= n!e^{-[(y_1)+(y_1+y_2)+\ldots+(y_1+y_2+\ldots+y_n)]} \\ &= n!e^{-[ny_1+(n-1)y_2+\ldots+y_n]} \\ &= \prod_{i=1}^{n} ie^{-iy_i}, \qquad y_i > 0. \end{aligned}$$

(4.3) It is from the solution to (4.2) that $Y_1, Y_2, \ldots, Y_n$ are independent random variables (since the joint density function can be rewritten as a product of individual density functions), and one can read the density function of Y_n as follows:

$$f_{Y_n}(y_n) = ne^{-ny_n}, \qquad \text{for } y_n > 0.$$

□

5 (3 points)

Let $(X_1, X_2)'$ be two dimensional random vector whose characteristic function is given as follows:

$$\varphi_{X_1,X_2}(t_1,t_2) = e^{it_1-2t_1^2-t_2^2-t_1t_2},$$

where i is the imaginary unit.
(5.1) (2p) Is $(X_1, X_2)'$ a two dimensional normal random vector? If yes, specify the mean vector μ and the covariance matrix Λ. If no, specify the reason(s).
(5.2) (1p) Find the distribution of $X_1 + X_2$. (Namely, specify which distribution with which parameters)

Solution. (5.1) Let us first pretend that $(X_1, X_2)'$ is a two dimensional normal random vector, then we need to find the mean vector μ and the covariance matrix Λ so that the characteristic function is

$$\varphi_{X_1,X_2}(t_1,t_2) = e^{i(t_1,t_2)\mu-\frac{1}{2}(t_1,t_2)\Lambda(t_1,t_2)'}.$$

By comparing this with $e^{it_1-2t_1^2-t_2^2-t_1t_2}$, we can easily obtain $\mu = (1, 0)'$. Now we try to find Λ so that

$$\frac{1}{2}(t_1,t_2)\Lambda(t_1,t_2)' = 2t_1^2 + t_2^2 + t_1t_2.$$

To this end, let $\Lambda = (a_{ij})_{1\le i,j\le 2}$. Then it holds that

$$\frac{1}{2}a_{11}t_1^2 + a_{12}t_1t_2 + \frac{1}{2}a_{22}t_2^2 = 2t_1^2 + t_2^2 + t_1t_2 \Longrightarrow a_{11} = 4, \quad a_{12} = 1, \quad a_{22} = 2.$$

Therefore, such $\mu = (1, 0)'$ and $\Lambda = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$ DO exist, implying that $(X_1, X_2)'$ is indeed a two dimensional normal random vector

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N(\mu, \Lambda).$$

(5.2) <u>Method 1</u>: One can obtain the characteristic function of $X_1 + X_2$ as follows

$$\varphi_{X_1+X_2}(t) = E(e^{it(X_1+X_2)}) = E(e^{itX_1+tX_2}) = \varphi_{X_1,X_2}(t,t) = e^{it-2t^2-t^2-t^2} = e^{it-4t^2} = \varphi_{N(1,8)}(t).$$

Therefore $X_1 + X_2 \sim N(1, 8)$.

<u>Method 2</u>: Since $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N(\mu, \Lambda)$, and $X_1 + X_2 = A\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ with $A = (1, 1)$, it follows that $X_1 + X_2$ is also normal with

$$\text{mean vector} = A\mu = 1, \qquad \text{covariance matrix} = A\Lambda A' = 8 \text{ (which is variance in this case)}.$$

So $X_1 + X_2 \sim N(1, 8)$.

□

6 (3 points)

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables with a common distribution function $F(x)$ (which is $F(x) = P(X_i \leq x)$). Let $F_n(x)$ be the empirical distribution function defined as

$$F_n(x) = \frac{\#\text{ observations among } X_1, X_2, \ldots, X_n \leq x}{n}.$$

For example, if we have observed $\{2,3,5,4\}$ for $\{X_1, X_2, X_3, X_4\}$ then $F_4(2.5) = \frac{1}{4}$ and $F_4(3.2) = \frac{2}{4}$.
(6.1) (1p) For each fixed x, prove that $F_n(x)$ converge to $F(x)$ in probability as $n \to \infty$.
(6.2) (2p) For each fixed x, determine $a(x)$ and $b(x)$, and show the following convergence in distribution

$$\frac{F_n(x) - a(x)}{b(x)/\sqrt{n}} \xrightarrow{d} N(0,1), \quad \text{as } n \to \infty.$$

Solution. (6.1) It is important to rewrite $F_n(x)$ as $F_n(x) = \frac{1}{n}(Y_1 + Y_2 + \ldots + Y_n)$ where $Y_i, 1 \leq i \leq n$ are i.i.d. with

Y_i	0	1
$p(y)$	$1 - F(x)$	$F(x)$

That is, if $X_i \leq x$, then $Y_i = 1$ and the corresponding probability is $P(X_i \leq x) = F(X)$. For Y_i,

$$\mu_Y = F(x), \qquad \sigma_Y^2 = F(x) - F(x)^2.$$

Weak law of large numbers (Theorem 6.5.1) implies that $F_n(x)$ converge to $\mu_Y = F(x)$ in probability.
(6.2) The central limit theorem (Theorem 6.5.2) implies that

$$\frac{F_n(x) - \mu_Y}{\sigma_Y/\sqrt{n}} \xrightarrow{d} N(0,1).$$

Therefore $a(x) = \mu_Y = F(x)$ and $b(x) = \sigma_Y = \sqrt{F(x) - F(x)^2}$. □

Discrete Distributions

Followingis a list of discrete distributions, abbreviations, their probability functions, means, variances, and characteristic functions. An asterisk (*) indicates that the expression is too complicated to present here; in some cases a closed formula does not even exist.

Distribution, notation	Probability function	$E\,X$	$\operatorname{Var} X$	$\varphi_X(t)$
One point $\delta(a)$	$p(a)=1$	a	0	e^{ita}
Symmetric Bernoulli	$p(-1)=p(1)=\frac{1}{2}$	0	1	$\cos t$
Bernoulli $\mathrm{Be}(p)$, $0\le p\le 1$	$p(0)=q,\ p(1)=p;\ q=1-p$	p	pq	$q+pe^{it}$
Binomial $\mathrm{Bin}(n,p)$, $n=1,2,\ldots$, $0\le p\le 1$	$p(k)=\binom{n}{k}p^kq^{n-k},\ k=0,1,\ldots,n;\ q=1-p$	np	npq	$(q+pe^{it})^n$
Geometric $\mathrm{Ge}(p)$, $0\le p\le 1$	$p(k)=pq^k,\ k=0,1,2,\ldots;\ q=1-p$	$\frac{q}{p}$	$\frac{q}{p^2}$	$\frac{p}{1-qe^{it}}$
First success $\mathrm{Fs}(p)$, $0\le p\le 1$	$p(k)=pq^{k-1},\ k=1,2,\ldots;\ q=1-p$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^{it}}{1-qe^{it}}$
Negative binomial $\mathrm{NBin}(n,p)$, $n=1,2,3,\ldots$, $0\le p\le 1$	$p(k)=\binom{n+k-1}{k}p^nq^k,\ k=0,1,2,\ldots;$ $q=1-p$	$n\frac{q}{p}$	$n\frac{q}{p^2}$	$(\frac{p}{1-qe^{it}})^n$
Poisson $\mathrm{Po}(m)$, $m>0$	$p(k)=e^{-m}\frac{m^k}{k!},\ k=0,1,2,\ldots$	m	m	$e^{m(e^{it}-1)}$
Hypergeometric $H(N,n,p)$, $n=0,1,\ldots,N$, $N=1,2,\ldots$, $p=0,\frac{1}{N},\frac{2}{N},\ldots,1$	$p(k)=\dfrac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}},\ k=0,1,\ldots,Np;$ $q=1-p;$ $n-k=0,\ldots,Nq$	np	$npq\frac{N-n}{N-1}$	*

Continuous Distributions

Following is a list of some continuous distributions, abbreviations, their densities, means, variances, and characteristic functions. An asterisk (*) indicates that the expression is too complicated to present here; in some cases a closed formula does not even exist.

Distribution, notation	Density	$E\,X$	$\operatorname{Var} X$	$\varphi_X(t)$
Uniform/Rectangular				
$U(a,b)$	$f(x)=\frac{1}{b-a},\ a<x<b$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$	$\frac{e^{itb}-e^{ita}}{it(b-a)}$
$U(0,1)$	$f(x)=1,\ 0<x<1$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{e^{it}-1}{it}$
$U(-1,1)$	$f(x)=\frac{1}{2},\ \lvert x\rvert<1$	0	$\frac{1}{3}$	$\frac{\sin t}{t}$
Triangular				
$\mathrm{Tri}(a,b)$	$f(x)=\frac{2}{b-a}\left(1-\frac{2}{b-a}\left\lvert x-\frac{a+b}{2}\right\rvert\right)$, $a<x<b$	$\frac{1}{2}(a+b)$	$\frac{1}{24}(b-a)^2$	$\left(\frac{e^{itb/2}-e^{ita/2}}{\frac{1}{2}it(b-a)}\right)^2$
$\mathrm{Tri}(-1,1)$	$f(x)=1-\lvert x\rvert,\ \lvert x\rvert<1$	0	$\frac{1}{6}$	$\left(\frac{\sin\frac{t}{2}}{\frac{t}{2}}\right)^2$
Exponential $\mathrm{Exp}(a),\ a>0$	$f(x)=\frac{1}{a}e^{-x/a},\ x>0$	a	a^2	$\frac{1}{1-ait}$
Gamma $\Gamma(p,a),\ a>0,\ p>0$	$f(x)=\frac{1}{\Gamma(p)}x^{p-1}\frac{1}{a^p}e^{-x/a},\ x>0$	pa	pa^2	$\frac{1}{(1-ait)^p}$
Chi-square $\chi^2(n),\ n=1,2,3,\ldots$	$f(x)=\frac{1}{\Gamma(\frac{n}{2})}x^{\frac{1}{2}n-1}\left(\frac{1}{2}\right)^{n/2}e^{-x/2},\ x>0$	n	$2n$	$\frac{1}{(1-2it)^{n/2}}$
Laplace $L(a),\ a>0$	$f(x)=\frac{1}{2a}e^{-\lvert x\rvert/a},\ -\infty<x<\infty$	0	$2a^2$	$\frac{1}{1+a^2t^2}$
Beta $\beta(r,s),\ r,s>0$	$f(x)=\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}x^{r-1}(1-x)^{s-1}$, $0<x<1$	$\frac{r}{r+s}$	$\frac{rs}{(r+s)^2(r+s+1)}$	*

Continuous Distributions (continued)

Distribution, notation	Density	$E\,X$	$\operatorname{Var} X$	$\varphi_X(t)$
Weibull $W(\alpha,\beta)$, $\alpha,\beta>0$	$f(x)=\frac{1}{\alpha\beta}x^{(1/\beta)-1}\,e^{-x^{1/\beta}/\alpha},\ x>0$	$\alpha^{\beta}\,\Gamma(\beta+1)$	$a^{2\beta}\big(\Gamma(2\beta+1) -\Gamma(\beta+1)^2\big)$	*
Rayleigh $\mathrm{Ra}\,(\alpha)$, $\alpha>0$	$f(x)=\frac{2}{\alpha}\,x\,e^{-x^2/\alpha},\ x>0$	$\frac{1}{2}\sqrt{\pi\alpha}$	$\alpha(1-\frac{1}{4}\pi)$	*
Normal $N(\mu,\sigma^2)$, $-\infty<\mu<\infty$, $\sigma>0$	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{1}{2}(x-\mu)^2/\sigma^2},$ $-\infty<x<\infty$	μ	σ^2	$e^{i\mu t-\frac{1}{2}t^2\sigma^2}$
$N(0,1)$	$f(x)=\dfrac{1}{\sqrt{2\pi}}\,e^{-x^2/2},\ -\infty<x<\infty$	0	1	$e^{-t^2/2}$
Log-normal $LN(\mu,\sigma^2)$, $-\infty<\mu<\infty,\ \sigma>0$	$f(x)=\dfrac{1}{\sigma x\sqrt{2\pi}}\,e^{-\frac{1}{2}(\log x-\mu)^2/\sigma^2},\ x>0$	$e^{\mu+\frac{1}{2}\sigma^2}$	$e^{2\mu}\big(e^{2\sigma^2}-e^{\sigma^2}\big)$	*
(Student's) t $t(n)$, $n=1,2,\ldots$	$f(x)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\Gamma(\frac{n}{2})}\cdot d\frac{1}{(1+\frac{x^2}{n})^{(n+1)/2}},$ $-\infty<x<\infty$	0	$\dfrac{n}{n-2},\ n>2$	*
(Fisher's) F $F(m,n)$, $m,n=1,2,\ldots$	$f(x)=\frac{\Gamma(\frac{m+n}{2})(\frac{m}{n})^{m/2}}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\cdot\frac{x^{m/2-1}}{(1+\frac{mx}{n})^{(m+n)/2}},$ $x>0$	$\frac{n}{n-2},$ $n>2$	$\frac{n^2(m+2)}{m(n-2)(n-4)}-\Big(\frac{n}{n-2}\Big)^2,$ $n>4$	*

Continuous Distributions (continued)

Distribution, notation	Density	$E\,X$	$\operatorname{Var} X$	$\varphi_X(t)$
Cauchy				
$C(m,a)$	$f(x) = \frac{1}{\pi} \cdot \frac{a}{a^2+(x-m)^2},\ -\infty < x < \infty$	$\nexists$	$\nexists$	$e^{imt-a\lvert t\rvert}$
$C(0,1)$	$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2},\ -\infty < x < \infty$	$\nexists$	$\nexists$	$e^{-\lvert t\rvert}$
Pareto	$f(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}},\ x > k$	$\frac{\alpha k}{\alpha - 1},\ \alpha > 1$	$\frac{\alpha k^2}{(\alpha-2)(\alpha-1)^2},\ \alpha > 2,$	*
$\mathrm{Pa}\,(k,\alpha),\ k > 0,\ \alpha > 0$				