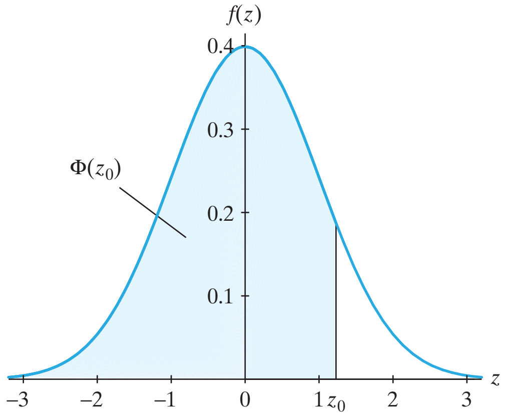
Probability and Statistics I

Mid-term Sample
SSE, CUHK(SZ)

Name: ______________________ Student ID: ______________________

Answer the multiple choice questions (Section I) in the Answer Card, and answer the regular questions (Section II) in the Answer Book.

Table Va The Standard Normal Distribution Function



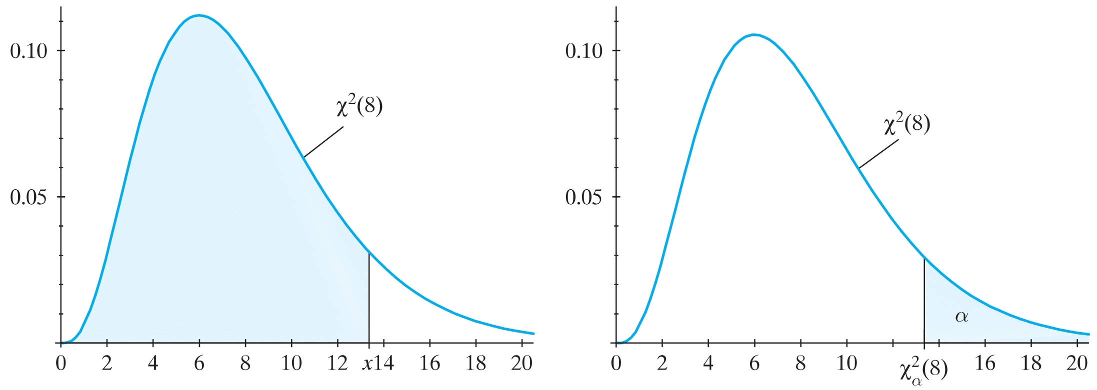
$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.326	2.576	2.807	3.291

Table IV The Chi-Square Distribution

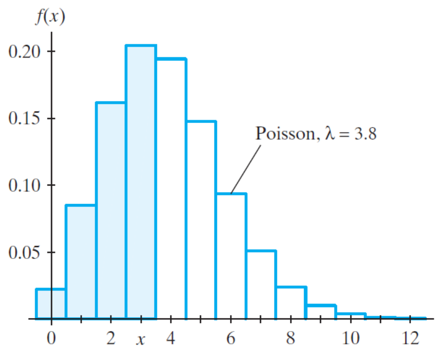


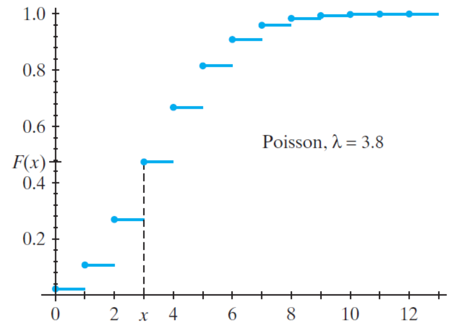
$$P(X \le x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2} dw$$

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58
16	5.812	6.908	7.962	9.312	23.54	26.30	28.84	32.00
17	6.408	7.564	8.672	10.08	24.77	27.59	30.19	33.41
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.80
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	35.56	38.88	41.92	45.64
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
40	22.16	24.43	26.51	29.05	51.80	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4
80	53.34	57.15	60.39	64.28	96.58	101.9	106.6	112.3

This table is abridged and adapted from Table III in *Biometrika Tables for Statisticians*, edited by E.S.Pearson and H.O.Hartley.

Table III The Poisson Distribution





$$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!}$$

	$\lambda = E(X)$									
x	4.2	4.4	4.6	4.8	5.0	5.2	5.4	5.6	5.8	6.0
0	0.015	0.012	0.010	0.008	0.007	0.006	0.005	0.004	0.003	0.002
1	0.078	0.066	0.056	0.048	0.040	0.034	0.029	0.024	0.021	0.017
2	0.210	0.185	0.163	0.143	0.125	0.109	0.095	0.082	0.072	0.062
3	0.395	0.359	0.326	0.294	0.265	0.238	0.213	0.191	0.170	0.151
4	0.590	0.551	0.513	0.476	0.440	0.406	0.373	0.342	0.313	0.285
5	0.753	0.720	0.686	0.651	0.616	0.581	0.546	0.512	0.478	0.446
6	0.867	0.844	0.818	0.791	0.762	0.732	0.702	0.670	0.638	0.606
7	0.936	0.921	0.905	0.887	0.867	0.845	0.822	0.797	0.771	0.744
8	0.972	0.964	0.955	0.944	0.932	0.918	0.903	0.886	0.867	0.847
9	0.989	0.985	0.980	0.975	0.968	0.960	0.951	0.941	0.929	0.916
10	0.996	0.994	0.992	0.990	0.986	0.982	0.977	0.972	0.965	0.957
11	0.999	0.998	0.997	0.996	0.995	0.993	0.990	0.988	0.984	0.980
12	1.000	0.999	0.999	0.999	0.998	0.997	0.996	0.995	0.993	0.991
13	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.996
14	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.999
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0
0	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.011	0.007	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000
2	0.043	0.030	0.020	0.014	0.009	0.006	0.004	0.003	0.002	0.001
3	0.112	0.082	0.059	0.042	0.030	0.021	0.015	0.010	0.007	0.005
4	0.224	0.173	0.132	0.100	0.074	0.055	0.040	0.029	0.021	0.015
5	0.369	0.301	0.241	0.191	0.150	0.116	0.089	0.067	0.050	0.038
6	0.527	0.450	0.378	0.313	0.256	0.207	0.165	0.130	0.102	0.079
7	0.673	0.599	0.525	0.453	0.386	0.324	0.269	0.220	0.179	0.143
8	0.792	0.729	0.662	0.593	0.523	0.456	0.392	0.333	0.279	0.232
9	0.877	0.830	0.776	0.717	0.653	0.587	0.522	0.458	0.397	0.341
10	0.933	0.901	0.862	0.816	0.763	0.706	0.645	0.583	0.521	0.460
11	0.966	0.947	0.921	0.888	0.849	0.803	0.752	0.697	0.639	0.579
12	0.984	0.973	0.957	0.936	0.909	0.876	0.836	0.792	0.742	0.689
13	0.993	0.987	0.978	0.966	0.949	0.926	0.898	0.864	0.825	0.781
14	0.997	0.994	0.990	0.983	0.973	0.959	0.940	0.917	0.888	0.854

I Multiple Choices (72 points)

- **3 points for each correct answer; -1 point for each incorrect answer; 0 points for no answer**
- **For each question, only choose (at most) one out of four given choices (A,B,C and D). If you choose more than one choice in one question, your answer will be incorrect and 1 point will be deduced.**

1. From a group of 3 first-year students, 4 sophomores, 4 juniors, and 3 seniors, a committee of size 4 is randomly selected. Find the probability that the committee will consist of (a) 1 from each class; (b) 2 sophomores and 2 juniors; (c) only sophomores or juniors.

   A. (a) 0.1439, (b) 0.0360, (c) 0.0699
   B. (a) 0.2510, (b) 0.0310, (c) 0.0952
   C. (a) 0.1539, (b) 0.4601, (c) 0.6905
   D. (a) 0.0438, (b) 0.3646, (c) 0.0799

   Solution: (A)
   (a) $\frac{(3\cdot4\cdot4\cdot3)}{\binom{14}{4}} = 0.1439$
   (b) $\frac{\binom{4}{2}\binom{4}{2}}{\binom{14}{4}} = 0.0360$
   (c)$\frac{\binom{8}{4}}{\binom{14}{4}} = 0.0699$

2. You ask your neighbor to water a sick plant while you are on vacation. Without water, it will die with probability 0.8; with water, it will die with probability 0.15. You are 90 percent certain that your neighbor will remember to water the plant. If the plant is dead upon your return, what is the probability that your neighbor forgot to water it?
   A. 0.372 B. 0.420 C. 0.101 D. 0.265

   Solution: (A)
   Let$A$ denote the event that the plant is alive and let W be the event that it was watered.

$P(A) = P(A \mid W)P(W) + P(A \mid W')P(W')$
$= (0.85)(0.9) + (0.2)(0.1) = 0.785$
$P(W' \mid A') = \frac{P(A'|W')P(W')}{P(A')} = \frac{(0.8)(0.1)}{1-0.785} = \frac{0.08}{0.215} = 0.372$

3. The random variable X has the probability density function

$$f(x) = \begin{cases} ax + bx^2, & 0 < x < 1, \\ 0, & \text{otherwise} \end{cases}$$

If $E(X) = 0.6$, find (a)$P(X < \frac{1}{2})$ and (b) $\text{Var}(X)$.

A. $P(X < \frac{1}{2}) = 0.4$, $\text{Var}(X) = 0.06$
B. $P(X < \frac{1}{2}) = 0.35$, $\text{Var}(X) = 0.06$
C. $P(X < \frac{1}{2}) = 0.35$, $\text{Var}(X) = 0.09$
D. $P(X < \frac{1}{2}) = 0.4$, $\text{Var}(X) = 0.09$

Solution: (B)
Since $1 = \int_0^1 (ax + bx^2)dx = \frac{a}{2} + \frac{b}{3}$
$0.6 = \int_0^1 (ax^2 + bx^3)dx = \frac{a}{3} + \frac{b}{4}$
we obtain $a = 3.6, b = -2.4$.Hence,
(a) $P(X < \frac{1}{2}) = \int_0^{1/2} (3.6x - 2.4x^2)dx = (1.8x^2 - 0.8x^3) \mid_0^{1/2} = 0.35$
(b) $E[X^2] = \int_0^1 (3.6x^3 - 2.4x^4)dx = 0.42$, so $\text{Var}(X) = 0.42 - 0.36 = 0.06$.

4. Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\theta = 10$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait
(a) more than 10 minutes;
(b) between 10 and 20 minutes.

A. (a) e, (b) $e^2 - e$
B. (a) e^{-10}, (b) $e^{-10} - e^{-20}$
C. (a) e, (b) $e^{-1} - e^{-2}$
D. (a) e^{-1}, (b) $e^{-1} - e^{-2}$

Solution: (D)
Let X denote the length of the call made by the person in the booth. $F(x)$ is the cdf at x in the solution. Then the desired probabilities are
(a) $P(X > 10) = 1 - F(10) = 1 - \int_0^{10} \frac{1}{10} e^{-\frac{x}{10}} dx = 1 - (-e^{-\frac{x}{10}}) \mid_0^{10} = e^{-1}$.
(b) $P(10 < X < 20) = F(20) - F(10) = e^{-1} - e^{-2}$

5. The following gambling game, known as the wheel of fortune (or chuck-a-luck), is quite popular at many carnivals and gambling casinos: A player bets on one of the numbers 1 through 6. Three dice are then rolled, and if the number bet by the player appears i times, $i = 1, 2, 3$, then the player wins i units; if the number bet by the player does not appear on any of the dice, then the player loses 1 unit. Let X denote the player's winnings in the game. (Actually, the game is played by spinning a wheel that comes to rest on a slot labeled by three of the numbers 1 through 6, but this variant is mathematically equivalent to the dice version.) Consider the following five statements:
(i) The random variable X follows binomial distribution.
(ii) $E(X) = \frac{17}{216}$
(iii) $P(X = 2) = \frac{15}{216}$
(iv) $P(X = 3) = \frac{1}{216}$
(v) This is an unfair game for the player. That is, $E(X) < 0$.

A. Only (i), (ii), (iii) and (iv) are true.
B. Only (iii), (iv) and (v) are true.
C. Only (ii), (iii) and (iv) are true.
D. Only (i), (iv) and (v) are true.

Solution: (B)
If we assume that the dice are fair and act independently of one another, then the number of times that the number bet appears is a binomial Hence, letting X denote the player's winnings in the game, we have $P(X = -1) = \binom{3}{0}(\frac{1}{6})^0(\frac{5}{6})^3 = \frac{125}{216}$
$P(X = 1) = \binom{3}{1}(\frac{1}{6})^1(\frac{5}{6})^2 = \frac{75}{216}$
$P(X = 2) = \binom{3}{2}(\frac{1}{6})^2(\frac{5}{6})^1 = \frac{15}{216}$

$P(X=3) = \binom{3}{3}(\frac{1}{6})^3(\frac{5}{6})^0 = \frac{1}{216}$
In order to determine whether or not this is a fair game for the player, let us calculate $E[X]$. From the preceding probabilities, we obtain
$E(X) = \frac{-125+75+30+3}{216} = \frac{-17}{216}$
Hence, in the long run, the player will lose 17 units per every 216 games he plays.

6. Find $P(X=4)$ if X has a Poisson distribution such that $3P(X = 1) = P(X=2)$.

   A. 0.285 B. 0.313 C. 0.238 D. 0.134

Solution: (D)
$3\frac{\lambda^1 e^{-\lambda}}{1!} = \frac{\lambda^2 e^{-\lambda}}{2!}$
$e^{-\lambda}\lambda(\lambda-6) = 0$
$\lambda = 6$
Thus $P(X=4) = P(X \leq 4) - P(X \leq 3) = 0.285 - 0.151 = 0.134$

7. Consider the moment generating function of X

$$M_X(t) = e^{3t+9t^2}.$$

   And Y has the moment generating function, for $t < 3$,

$$M_Y(t) = \frac{3e^{2t}}{(3-t)}.$$

   Statement 1 : Y has $E[Y]$ of $\frac{7}{9}$
   Statement 2 : Y has $\text{Var}(Y)$ of $\frac{1}{3}$
   Statement 3 : X follows normal distribution $N(3,9)$

   How many statements are correct?

   A. 0 B. 1 C. 2 D. 3

Solution: (A)
$E[Y] = \frac{\partial}{\partial t} M_Y(t) = \frac{21e^{2t} - 6te^{2t}}{(3-t)^2}|_{t=0} = \frac{7}{3}$ Take the second derivative to get $E[Y^2]$ and then $Var(Y) = \frac{1}{9}$.

The moment generating function of a normal distribution is $M(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$, so that $X \sim N(3, 18)$.

8. CUHKSZ students sometimes delay laundry for a few days.
A busy student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability 2/3 and 2 days with probability 1/3. Let B be the number of days a busy student delays laundry. What is $E[B]$?

A. 3 B. 4 C. $\frac{3}{4}$ D. $\frac{4}{3}$

Solution: (B)
The expected time to complete a problem set is:

$1 \times \frac{2}{3} + 2 \times \frac{1}{3} = \frac{4}{3}$ Therefore, the expected time to complete all three problem sets is: $E[B] = E[pset1] + E[pset2] + E[pset3] = \frac{4}{3} + \frac{4}{3} + \frac{4}{3} = 4$

9. CUHKSZ students sometimes delay laundry for a few days.
A relaxed student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $E[R]$?

A. 5 B. 6 C. $\frac{5}{6}$ D. $\frac{6}{5}$

Solution: (A)
If we regard doing laundry as a failure, then the mean time to failure is $1/(1/6) = 6$. However, this counts the day laundry is done, so the number of days delay is 6-1 = 5. Alternatively, we could derive the answer as follows:

$E[R] = \sum_{k=0}^{\infty} P(R > k) = \frac{5}{6} + (\frac{5}{6})^2 + (\frac{5}{6})^3 + ... = \frac{5}{6} \times (1 + \frac{5}{6} + (\frac{5}{6})^2 + ...) = \frac{5}{6} \times \frac{1}{1-\frac{5}{6}} = 5$

10. Suppose that Pandora Restaurant everyday (24 hours) receives 720 complaints on average. It is assumed that the number of complaints received follow an approximate Poisson process. Suppose that X (in month) is the time it takes for Pandora to reply to a complaint, and X follows a gamma distribution with mean $\frac{3}{2}$ months and standard deviation $\sqrt{\frac{3}{4}}$ months.

    Statement 1 : The probability that Pandora will have to wait longer than 21.06 minutes for the first 7th complain is 0.05.
    Statement 2 : X follows the gamma distribution with $\alpha = 4$.
    Statement 3 : Let $Y = 3X$. Then Y has a gamma distribution with the moment generating function $(\frac{\frac{2}{3}}{\frac{2}{3} - t})^3$

    Which statement is true?

    A. Statements 1 and 3 only
    B. Statements 2 and 3 only
    C. Statements 3 only
    D. None of them

Solution: (C)
1: The mean rate of complains per minute is $\lambda = \frac{1}{2}$. Thus $\theta = 2$ and $\alpha = \frac{r}{2} = 7$. If Z denotes the waiting time until the 7th complain, then Z is $\chi^2(14)$. $P(Z > 21.06) = 0.1$.
2: $\mu = \alpha\theta = \frac{3}{2}$, $\sigma^2 = \alpha\theta^2 = \frac{3}{4}$. $\alpha = 3$.
3: $M_X(t) = (\frac{2}{2-t})^3$, $M_Y(t) = E[e^{tY}] = E[^{t3X}] = M_X(3t) = (\frac{2}{2-3t})^3$

11. Consider these 3 statements. We denote A' and B' to be the complement of the set A and B respectively.

Statement 1: If events A and B are mutually exclusive and exhaustive, A' and B' are mutually exclusive.
Statement 2: If events A and B are mutually exclusive but not exhaustive, A' and B' are exhaustive.
Statement 3: If events A and B are exhaustive but not mutually exclusive, A' and B' are exhaustive.

Choose the correct option.

A. Statement1–False, Statement2–True, Statement3–True
B. Statement1–True, Statement2–True, Statement3–False
C. Statement1–True, Statement2–False, Statement3–False
D. Statement1–False, Statement2–True, Statement3–False

Solution: (B)
S1: $A' \cap B' = (A \cup B)' = S' = \emptyset$. Thus the events A' and B' are mutually exclusive.
S2: Let $C = (A' \cup B')'$, that is the part that is not contained in $A' \cup B'$. Using De Morgan's Law $C = A \cap B = \emptyset$. Thus, there is nothing that is not a part of A' or B'. Hence, A' and B' are mutually exhaustive.
S3: As in previous part, let $C = (A' \cup B')' = A \cap B$ which is not null. Thus, A' and B' are not mutually exhaustive.

12. Suppose we roll two fair six-sided dice, one red and one blue. Let A be the event that the two dice show the same value. Let B be the event that the sum of the two dice is equal to 12. Let C be the event that the red die shows 4. Let D be the event that the blue die shows 4. Consider the following 5 statements.

    1. A and B are independent.
    2. A and C are independent.
    3. A and D are independent.
    4. C and D are independent.
    5. A, C and D are mutually independent.

    Which statement is true?

A. 2 only
B. 2 and 3 only
C. 2 and 3 and 4 only
D. All of them

Solution: (C)

13. Consider following 4 statements:

(i) If X follows a Gamma distribution with parameter $\alpha = 2$ and $\theta = 1$, then $E[X^4] = 120$
(ii) Consider the Gamma function with a positive integer α, then we could write it as $\Gamma(\alpha) = (\alpha - 1)!$
(iii) If $X \sim N(3, 1)$, then $E[X^3]$=27
(iv) If X follows a Chi-square distribution with 1 degree of freedom, then $P(X \geq 2.706) = 0.9$

A. All statements are ture.
B. Only (i), (ii) and (iii) are ture.
C. Only (iii) and (iv) are ture.
D. Only (i) and (ii) are ture.

Solution: (D)
(i). Using the moment generating function of gamma distribution, we have $E[X^4]$=$\alpha(\alpha + 1)(\alpha + 2)(\alpha + 3)\theta^4 = 120$.
(ii). This a property of Gamma function and it can be proved by using integration by part.
(iii). Let $X = Z + 3$, where $Z \sim N(0, 1)$. Notice that $E[Z] = 0$, $E[Z^2] = 1$ and $E[Z^3] = 0$, we have

$$\mathbb{E}[X^3] = \mathbb{E}[Z^3 + 9Z^2 + 27Z + 27] = 36$$

(iv). This can be checked through the table. The correct one is $P(X \leq 2.706) = 0.9$
Therefore, only (i) and (ii) are true.

14. The weekly amount of downtime X (in hours) for a certain industrial machine has following probability density function

$$f(x) = \begin{cases} \frac{1}{16}x^2e^{-x/2}, & x \geq 0, \\ 0, & \text{otherwise} \end{cases}$$

Consider the following statements:

(i) The random variable X follows a exponential distribution.
(ii) $E(X) = 1/2$, Var(X)=1/4
(iii) The 10th percentile of the random variable X is 2.204.
(iv) The inequality $P(X \geq 2.204) < \frac{E(X)}{2.204}$ holds.

A. Only (i) and (ii) are true
B. Only (i) and (iii) are true
C. Only (ii) and (iv) are true
D. Only (iii) and (iv) are true

Solution: (D)
From the expression of probability density function we find that X actually follows a Gamma distribution with parameters $\alpha = 3$, $\theta = 2$, and equivalently X follows a Chi-square distribution with degree of freedom $r = 6$ (since $\theta = 2$ and $\alpha = 3$). Therefore,
(i) is wrong since $\alpha \neq 1$.
(ii) is wrong, since $E(X) = \alpha\theta = 6$, $Var(X) = \alpha\theta^2 = 12$.
(iii) is correct. The 10th percentile for the random variable X is a number such that $P\left(X \leq \chi^2_{0.9}(6)\right) = 0.1$, where $\alpha = 0.1$ in this case. The value of $\chi^2_{0.9}(6)$ could be obtained by checking the table of Chi-square distribution.
(iv) is correct. We find that $P(X \geq 2.204) = 1 - 0.1 = 0.9$ follows from (iii), and since $E(X) = 6$, the inequality holds(the inequality is known as the Markov inequality but we actually do not need to know it to find the answer).

15. Given a random variable X with following probability density function

$$f(x) = \begin{cases} 4x^3, & 0 < x < 1, \\ 0, & \text{otherwise} \end{cases}$$

What's the median of this distribution?

A. 2^{-1} B. 2^{-2} C. $2^{-1/2}$ D. $2^{-1/4}$

Solution: (D)
Let m be the value of median. We have $\int_0^m 4x^3\ dx = 1/2$, and yields that $m = 2^{-1/4}$.

16. Suppose X follows a uniform distribution on the interval $[-2, 1]$, that is, $X \sim U[-2, 1]$. Let $Y = X^2$. Which one of the following is the cumulative distribution function of Y?

A. $F_Y(y) = \begin{cases} 2\sqrt{y}/3, & 0 \leq y \leq 1, \\ \left(1+\sqrt{y}\right)/3, & 1 < y \leq 4, \\ 0, & y < 0, \\ 1, & y > 4 \end{cases}$

B. $F_Y(y) = \begin{cases} \left(1+\sqrt{y}\right)/3, & 0 \leq y \leq 1 \\ \sqrt{y}/2, & 1 < y \leq 4 \\ 0, & y < 0, \\ 1, & y > 4 \end{cases}$

C. $F_Y(y) = \begin{cases} (1+\sqrt{y})/3, & 0 \leq y \leq 4 \\ 0, & y < 0, \\ 1, & y > 4 \end{cases}$

D. $F_Y(y) = \begin{cases} 2\sqrt{y}/3, & 0 \leq y \leq 4 \\ 0, & y < 0, \\ 1, & y > 4 \end{cases}$

Solution: (A)
The cdf of X is given by $F_X(x) = (x+2)/3$, where $-2 \leq x \leq 1$
Consider the cdf of Y for $0 \leq y \leq 1$ (therefore $-1 \leq x \leq 1$), we have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \end{aligned}$$

$$= P(X \le \sqrt{y}) - P(X \le -\sqrt{y})$$
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$
$$= 2\sqrt{y}/3$$

Then, we consider the case $1 < y \le 4$, therefore $-2 \le x \le -1$ and we have

$$\begin{aligned} F_Y(y) = P(Y \le y) &= P(Y < 1) + P(1 \le Y \le y) \\ &= \frac{2}{3} + P(1 \le X^2 \le y) \\ &= \frac{2}{3} + P(-\sqrt{y} \le X \le -1) \\ &= \frac{2}{3} + P(X \le -1) - P(X \le -\sqrt{y}) \\ &= \frac{2}{3} + F_X(-1) - F_X(-\sqrt{y}) \\ &= \left(1 + \sqrt{y}\right)/3 \end{aligned}$$

Combine them we have

$$F_Y(y) = \begin{cases} 2\sqrt{y}/3, & 0 \le y \le 1, \\ \left(1 + \sqrt{y}\right)/3, & 1 < y \le 4, \\ 0, & y < 0, \\ 1, & y > 4 \end{cases}$$

17. Suppose there are two well defined events, event A and event B. The probability that only one of them occurs is 0.3, and $P(A)+P(B) = 0.5$. What's the probability that at least one of them not occur?
A. 0.3 B. 0.5 C. 0.6 D. 0.9

Solution: (D)
"Only one of them occurs" means, $P\left((A \cap B') \cup (A' \cap B)\right) = 0.3$. That is,

$$\begin{aligned} 0.3 &= P\left((A \cap B') \cup (A' \cap B)\right) \\ &= P(A) - P(A \cap B) + P(B) - P(A \cap B) \\ &= 0.5 - 2P(A \cap B) \end{aligned}$$

Therefore,

$$P(A \cap B) = 0.1$$
$$P(A' \cup B') = 1 - P(A \cap B) = 0.9.$$

18. Let X be a random variable with following probability mass function

$$f(x) = \begin{cases} \frac{c}{x!}, & \text{for } x\text{=0,1,2,}\ldots, \\ 0, & \text{otherwise} \end{cases}$$

where c is a constant.
Determine the value of c and compute the expectation of X.
A. $c = e^{-2},\ E[X] = e^{-1}$
B. $c = e^{-1},\ E[X] = 1$
C. $c = 1,\ E[X] = e$
D. $c = e,\ E[X] = e^2$

Solution: (B)
By the power series of e and the properties of probability mass function, we have

$$\sum_{x=0}^{\infty} \frac{c}{x!} = c(\sum_{x=0}^{\infty} \frac{1}{x!}) = ce = 1.$$

Therefore $c = 1/e$, and $E[X]$ is given by

$$E[X] = \sum_{x=0}^{\infty} x\frac{c}{x!} = c\sum_{x=1}^{\infty} \frac{1}{(x-1)!} = ce = 1.$$

19. Suppose $P(A|B) = P(B|A) = \frac{1}{4}, P(A') = \frac{2}{3}$, Which one of the following claims is true about these statements?
A. A and B are independent, and $P(A \cap B) = \frac{5}{12}$.
B. A and B are independent, and $P(A) = P(B)$.

C. A and B are dependent, and $P(A \cup B) = \frac{7}{12}$.
D. A and B are dependent, and $P(A|B') = P(A|B)$

Solution: (C)
Since
$$P(A|B) = P(B|A) = \frac{1}{4}$$
That is,
$$\frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(A)} = \frac{1}{4}.$$
Therefore
$$P(A \cap B) \neq P(A)P(B)$$
Then,
$$P(A) = 1 - P(A') = \frac{1}{3}, P(B) = \frac{1}{3}, P(A \cap B) = \frac{1}{4}P(A) = \frac{1}{12}$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{3} + \frac{1}{3} - \frac{1}{12} = \frac{7}{12}.$$

20. Products produced by a machine has a 3% defective rate. The first 10 inspections have been found to be free of defectives. What is the probability that the first defective will occur on the 15th inspection?
A. 0.97^4
B. $0.97^5 \times 0.03$
C. $1 - 0.97^4 \times 0.03$
D. $0.97^4 \times 0.03$

Solution: (D)
Let X be the number of products need to be checked to detect the first defective product. X has geometric distribution with $p = 0.03, q = 0.97$.

$$P(X = 15|X > 10) = \frac{P(X = 15 \cap X > 10)}{P(X > 10)}$$

$$= \frac{P(X = 15)}{P(X > 10)}$$
$$= \frac{q^{14}p}{\sum_{k=11}^{\infty} q^{k-1}p} = \frac{q^{14}p}{q^{10}} = q^4 p$$

Since $q = 0.97$. Therefore it should be D.

21. Suppose that $X \sim N(3, 2^2)$, find the largest d such that $P(X > d) \geq 0.9$.
A. 5.56 B. 4.68 C. 2.08 D. 0.44

Solution: (D)

$$P(X > d) = 1 - P(X \leq d) = 1 - P(\frac{X-3}{2} \leq \frac{d-3}{2})$$
$$= 1 - \Phi(\frac{d-3}{2}) \geq 0.9$$
$$\Phi(\frac{d-3}{2}) \leq 0.1$$

Check that table we know that $\Phi(1.28) = 0.9$.And $\Phi(-1.28) = 1 - \Phi(1.28) = 0.1$ Therefore $\frac{d-3}{2} = -1.28$, that is $d \leq 0.44$.

22. Suppose that the moment-generating function $M_X(t)$ of the continuous random variable X has the property $M_X(t) = e^t M_X(-t)$ for all t. What is $E(X)$?
A. $\frac{1}{4}$ B. $\frac{1}{2}$ C. 2 D. 4

Solution: (B)
Since $M_X(t) = e^t M_X(-t)$ can be written as $E(e^{tX}) = e^t E(e^{-tX}) = E(e^{t(1-X)})$ and thus $M_X(t) = M_{1-X}(t)$ for all t. Since the moment-generating function determines uniquely the probability distribution, it follows that the random variable X has the same distribu-

tion as the random variable $1 - X$. Hence $E(X) = E(1 - X)$ and so $E(X) = \frac{1}{2}$.

23. Let X be lifetime (measured in hours) of a cetain type of electronic device, and its probabiltiy density function given by

$$f(x) = \begin{cases} \frac{10}{x^2}, & x > 10 \\ 0, & x \leq 10 \end{cases}$$

What is the probability that at least 2 of 4 such types of devices will function for at least 15 hours?

A. $\frac{57}{81}$ B. $\frac{59}{81}$ C. $\frac{69}{81}$ D. $\frac{72}{81}$

Solution: (D)

Let X be the lifetime (measured in hours) of a certain type of device. Then,

$$\begin{aligned} P(X \geq 15) &= \int_{15}^{\infty} \frac{10}{x^2} dx = [-\frac{10}{x}]_{15}^{\infty} \\ &= \lim_{x \to \infty} -\frac{10}{x} - (-\frac{10}{15}) \\ &= -\lim_{x \to \infty} \frac{10}{x} + \frac{2}{3} = \frac{2}{3} \end{aligned}$$

Let Y be the number of devices that function at least 15 hours. Then,

$$P(X \geq 2) = 1 - \binom{4}{0}(\frac{2}{3})^0(\frac{1}{3})^4 - \binom{4}{1}(\frac{2}{3})^1(\frac{1}{3})^3 = \frac{72}{81}$$

24. Consider 3 urns. Urn A contains 2 white and 4 red balls; urn B contains 8 white and 4 red balls; and urn C contains 1 white and 3 red balls. If 1 ball is selected from each urn, what is the probability that the ball chosen from urn A was white, given that exactly 2 white balls were selected?

A. $\frac{6}{11}$ B. $\frac{9}{11}$ C. $\frac{8}{11}$ D. $\frac{7}{11}$

Solution: (D)
The probability in the question should be

$$P(\text{Ball from A white}|\text{2 white balls selected})$$
$$= \frac{P(\text{Ball from A white} \cap \text{2 white balls selected})}{P(\text{2 white balls selected})}$$

First, consider $P(\text{Ball from } A \text{ white} \cap \text{2 white balls selected})$. This means that the ball chosen from A must be white and either the ball from B or C is white and the other one is not. The probability of drawing a white ball from A is $\frac{2}{6}$. Likewise, the probability of drawing a white ball from B, C is $\frac{8}{12}$ and $\frac{1}{4}$ respectively. The probability of not drawing the white ball from B, C is $\frac{4}{12}$ and $\frac{3}{4}$ respectively. So,

$$P(\text{Ball from A white} \cap \text{2 white balls selected}) = \frac{2}{6}(\frac{8}{12} \cdot \frac{3}{4} + \frac{1}{4} \cdot \frac{4}{12})$$

Now consider $P(\text{2 white balls selected})$. There are 3 ways we can choose the two white balls: choose a white ball from A and B and a red ball from C, choose a white ball from B and C and a red ball from A and choose a white ball from A and C and a red ball from B. So the denominator is

$$\frac{2}{6} \cdot \frac{8}{12} \cdot \frac{3}{4} + \frac{4}{6} \cdot \frac{8}{12} \cdot \frac{1}{4} + \frac{2}{6} \cdot \frac{4}{12} \cdot \frac{1}{4}$$

Therefore the probability is $\frac{7}{11}$.

II Regular Questions (28 points)

25. (a) A random variable X with parameter $\theta \in \mathbb{R}$ belongs to the exponential family if its probability mass function or probability density function can be written as

$$f(x) = h(x)e^{\theta x - A(\theta)},$$

where $h(x)$ and $A(\theta)$ are some known functions. Note that $h(x)$ does not depend on θ and $A(\theta)$ does not depend on x.

(a1) (2 points) Show that a Poisson distribution with mean $\lambda > 0$

belongs to the exponential family by identifying appropriate $h(x)$, θ and $A(\theta)$.

(a2) (3 points) Assume that X is a continuous random variable that belongs to the exponential family. Show that the moment generating function $M(t)$ of X can be written as

$$M(t) = E(e^{tX}) = e^{A(\theta+t)-A(\theta)}.$$

(a3) (3 points) Continuing part (a2), and assume that $A(\theta)$ is twice differentiable so that its first and second derivative with respect to θ exist. Show that the mean and variance of X are

$$E(X) = \frac{d}{d\theta}A(\theta), \quad \mathrm{Var}(X) = \frac{d^2}{d\theta^2}A(\theta).$$

(b) Let n be a fixed positive integer, and X be a U(0, n) random variable, that is, a continuous uniform distribution on the interval $(0, n)$. Define $Y = \lfloor X \rfloor$, where for a real number x, $\lfloor x \rfloor$ is x rounded down to the nearest integer (e.g., $\lfloor 5.7 \rfloor = 5$).

(b1) (3 points) Find $f(y)$, the probability mass function (pmf) of Y.

(b2) (2 points) Find $F(y)$, the cumulative distribution function (cdf) of Y.

(b3) (1 points) Find $E(Y)$, the mean of Y.

Solution: Part (a1). The probability mass function of Poisson random variable with mean λ is given by

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{1}{x!}e^{(\ln\lambda)x}e^{-\lambda} = h(x)e^{\theta x - A(\theta)},$$

where we take

$$\begin{aligned} h(x) &= \frac{1}{x!} \\ \theta &= \ln\lambda \\ A(\theta) &= \lambda. \end{aligned}$$

Part (a2).

$$\begin{aligned}E(e^{tX}) &= \int e^{tx} h(x) e^{\theta x - A(\theta)}\, dx \\ &= e^{-A(\theta)} \int h(x) e^{(\theta+t)x}\, dx \\ &= e^{A(\theta+t)-A(\theta)} \int h(x) e^{(\theta+t)x - A(\theta+t)}\, dx \\ &= e^{A(\theta+t)-A(\theta)},\end{aligned}$$

where in the last equality we use $\int h(x)e^{(\theta+t)x-A(\theta+t)}\, dx = 1$.

Part (a3).

$$\frac{d}{dt}E(e^{tX}) = e^{A(\theta+t)-A(\theta)}\frac{d}{dt}A(\theta+t)$$
$$\frac{d^2}{dt^2}E(e^{tX}) = e^{A(\theta+t)-A(\theta)}\frac{d^2}{dt^2}A(\theta+t) + e^{A(\theta+t)-A(\theta)}\left(\frac{d}{dt}A(\theta+t)\right)^2.$$

As a result, by taking $t=0$ above, we have

$$\begin{aligned}E(X) &= \frac{d}{dt}E(e^{tX})\bigg|_{t=0} = \frac{d}{d\theta}A(\theta), \\ E(X^2) &= \frac{d^2}{dt^2}E(e^{tX})\bigg|_{t=0} = \frac{d^2}{d\theta^2}A(\theta) + \left(\frac{d}{d\theta}A(\theta)\right)^2, \\ \mathrm{Var}(X) &= E(X^2) - E(X)^2 = \frac{d^2}{d\theta^2}A(\theta).\end{aligned}$$

Part (b1). Y can take values in the set $\bar{S} = \{0, 1, 2, \ldots, n-1\}$. For $y \in \{0, 1, 2, \ldots, n-1\}$,

$$f(y) = P(Y=y) = P(y \le X < y+1) = \int_y^{y+1} \frac{1}{n}\, dx = \frac{1}{n}.$$

As a result, we have

$$f(y) = \begin{cases} \frac{1}{n}, & y \in \{0, 1, 2, \ldots, n-1\}, \\ 0, & \text{otherwise.}\end{cases}$$

In other words, Y has a discrete uniform distribution on $\bar{S}$.

Part (b2).

$$F(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{n}, & 0 \le y < 1, \\ \frac{2}{n}, & 1 \le y < 2, \\ \vdots & \vdots \\ \frac{n-1}{n}, & n-2 \le y < n-1, \\ 1, & n-1 \le y. \end{cases}$$

Part (b3).

$$E(Y) = \frac{1}{n}\sum_{y=1}^{n-1} y = \frac{n-1}{2}.$$

26. Consider a normal random variable $X \sim N(\mu, \sigma^2)$ with $\sigma > 0$.

(a1) (4 points) Show that for any $a < b$,

$$P(-b \le X \le -a) = P(a + 2\mu \le X \le b + 2\mu).$$

(a2) (8 points) **By using the moment generating function technique**, show that $(X^2 - 2\mu X + \mu^2)/\sigma^2$ has a Chi-square distribution with degrees of freedom 1, i.e.,

$$(X^2 - 2\mu X + \mu^2)/\sigma^2 \sim \chi^2(1)$$

(a3) (2 points) What is $E\left(\frac{X^2-2\mu X+\mu^2}{\sigma^2}\right)^2$?

Solution: Part (a). There are at least two possible solutions. One is based on the property of the cdf of $N(0,1)$, i.e., $\Phi(x)$ that $\Phi(-x) = 1 - \Phi(x)$. The other one is based on the coordinate change.

The first solution is

$$\begin{aligned}
P(-b \leq X \leq -a) &= P\left(\frac{-b-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{-a-\mu}{\sigma}\right) \\
&= \Phi\left(\frac{-a-\mu}{\sigma}\right) - \Phi\left(\frac{-b-\mu}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{a+\mu}{\sigma}\right) - 1 + \Phi\left(\frac{b+\mu}{\sigma}\right) \\
&= \Phi\left(\frac{b+\mu}{\sigma}\right) - \Phi\left(\frac{a+\mu}{\sigma}\right) \\
&= P\left(\frac{a+\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b+\mu}{\sigma}\right) \\
&= P(a+2\mu \leq X \leq b+2\mu)
\end{aligned}$$

Let $f(x)$ be the pdf of X. Then the second solution is simply to show that

$$\int_{-b}^{-a} f(x)dx = \int_{a+2\mu}^{b+2\mu} f(z)dz.$$

This is true by noting the pdf of X and taking the coordinate change

$$z = -x + 2\mu.$$

Part (b). The question is equivalent to show by moment generating function technique that if $Y \sim N(0,1)$, $Y^2 \sim \chi^2(1)$. By the definition of moment generating function, we have

$$M(t) = Ee^{tY^2} = \int_{-\infty}^{\infty} e^{ty^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(t-\frac{1}{2})y^2} dy$$

For $t < 1/2$, we further have

$$\begin{aligned}
M(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(1-2t)y^2} dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y}{\frac{1}{(1-2t)^{\frac{1}{2}}}}\right)^2} dy \\
&= \frac{1}{(1-2t)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{1}{(1-2t)^{\frac{1}{2}}}} e^{-\frac{1}{2}\left(\frac{y}{\frac{1}{(1-2t)^{\frac{1}{2}}}}\right)^2} dy
\end{aligned}$$

Since

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{1}{(1-2t)^{\frac{1}{2}}}} e^{-\frac{1}{2}\left(\frac{y}{\frac{1}{(1-2t)^{\frac{1}{2}}}}\right)^2} dy = 1$$

we have

$$M(t) = \frac{1}{(1-2t)^{\frac{1}{2}}}, \quad t < \frac{1}{2},$$

which is the moment generating function of $\chi^2(1)$. This completes the proof.

Part (c). Let $Z \sim \chi^2(1)$. Then the question is to calculate $E(Z^2)$. Since

$$E(Z) = 1, Var(Z) = 2,$$

$$E(Z^2) = Var(Z) + (E(Z))^2 = 2 + 1 = 3.$$

The same result can be obtained by calculating $M''(0) = 3$.