Tutorial Worksheet 1 - Reviews of Probability Theory

Problem 1. (Transformation of Random Variables)

(a) (a) Let X_1, X_2, X_3 be independent $N(0,1)$ random variables. Find the probability density function of $U = X_1^2 + X_2^2 + X_3^2$.

(b) Suppose that the random vector $\boldsymbol{Y} = (Y_1, Y_2, Y_3)$ is uniformly distributed on the sphere of radius 1 centred at the origin; that is, $\boldsymbol{Y}$ has joint probability density function (pdf)

$$f_{\boldsymbol{Y}}(y_1, y_2, y_3) = \begin{cases} \frac{3}{4\pi} & \text{if } (y_1, y_2, y_3) \in S \\ 0 & \text{otherwise,} \end{cases}$$

where $S = \left\{(y_1, y_2, y_3) \in \mathbb{R}^3 : y_1^2 + y_2^2 + y_3^2 \leq 1\right\}$ is the sphere of radius 1 centred at $(0,0,0)$. Let $V = Y_1^2 + Y_2^2 + Y_3^2$. Find the probability density function of V and find $E[V]$.

(c) Let X and Y be independent and identically distributed uniform random variables on the interval $(0,1)$. Define $U = \frac{1}{2}(X+Y)$ to be the average and define $V = X$.

(i) Find the joint probability density function of (U, V) and draw the sample space of (U, V).

(ii) Find the marginal probability density function of U.

(d) Let $X = (X_1, X_2)^T$ be uniformly distributed on the positive quadrant intersected with the disk of radius 1 centred at the origin; i.e., X has joint pdf $f_X(x_1, x_2) = \frac{4}{\pi} I_{S_X}(x_1, x_2)$, where

$$S_X = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1^2 + x_2^2 \leq 1\right\}.$$

Let $Y_1 = X_1^2$ and $Y_2 = X_2^2$. Find the joint pdf of $(Y_1, Y_2)^T$ and the marginal pdf of Y_1.

(e) Let X_1, X_2, X_3 be independent and identically distributed Exponential (λ) random variables (the Exponential (λ) distribution has pdf $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$ and $f_X(x) = 0$ for $x \leq 0$, and df $F_X(x) = 1 - e^{-\lambda x}$ for $x > 0$ and $F_X(x) = 0$ for $x \leq 0$). Find $P\left(X_1 + X_2 + X_3 \leq \frac{3}{2}\right)$. (Write down the appropriate 3-dimensional integral and evaluate it).

Problem 2. (Order Statistics)

(a) Let $X_1, \ldots, X_n$ be a sequence of independent random variables uniformly distributed on the interval $(0,1)$, and let $X_{(1)}, \ldots, X_{(n)}$ denote their order statistics. For fixed k, let $g_n(x)$ denote the probability density function of $nX_{(k)}$. Find $g_n(x)$ and show that

$$\lim_{n \to \infty} g_n(x) = \begin{cases} \frac{x^{k-1}}{(k-1)!} e^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

which is the Gamma$(k, 1)$ density.

(b) (i) Let $X_1, \ldots, X_6$ be independent random variables uniformly distributed on the interval $(0,1)$. Find the pdf of $U = \min\left\{\max\left(X_1, X_2\right), \max\left(X_3, X_4\right), \max\left(X_5, X_6\right)\right\}$.

(ii) Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, each with a Uniform distribution on the interval $(0,1)$. Let $X = \min\left(X_1, \ldots, X_n\right)$ and $Y = \max\left(X_1, \ldots, X_n\right)$. Find $P\left(X < \frac{1}{2} < Y\right)$ and $E\left[X^3\right]$.

(c) Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ be independent Uniform $(0,1)$ random variables. We form n rectangles, where the i-th rectangle has adjacent sides of length X_i and Y_i, for $i = 1, \ldots, n$. Let A_i be the area of the i th rectangle, $i = 1, \ldots, n$, and define $A_{\max} = \max\left(A_1, \ldots, A_n\right)$. Find the pdf of $A_{\max}$.

(d) Let $X_1, \ldots, X_n$ be independent and identically distributed Exponential (λ) random variables. Compute $E\left[X_{(1)} e^{-\lambda X_{(2)}}\right]$, where $X_{(1)}$ and $X_{(2)}$ are the first and second order statistics of $X_1, \ldots, X_n$.

(e) Let X_1, X_2, X_3 be independent Uniform $(0,1)$ random variables, and let $X_{(1)}, X_{(2)}, X_{(3)}$ denote their order statistics. Let X be the area of the square with side length $X_{(2)}$ and let Y be the area of the rectangle with side lengths $X_{(1)}$ and $X_{(3)}$. Find $P(X > Y)$, $E[X]$, and $E[Y]$.

Problem 3. (Convergence of Random Variables and Limit Theorems)

(a) Let X and $X_1, X_2, \ldots$ be random variables each having a $N(0,1)$ distribution. Suppose (X_n, X) has a bivariate normal distribution for each n and the correlation between X_n and X is $\rho(X_n, X) = \rho_n$, for $n \geq 1$.

- (i) Show that X_n converges to X in distribution as $n \to \infty$ (for arbitrary correlations ρ_n).
- (ii) If $\rho_n \to 1$ as $n \to \infty$, show that X_n converges to X in probability as $n \to \infty$.
- (iii) Show that if $\rho_n = 1 - a^n$ for some constant $a \in (0,1)$, then X_n converges to X with probability 1 as $n \to \infty$. Do you get convergence with probability 1 if $a = 0$? If $a = 1$? Prove your answers.

(b) Suppose 80 points are uniformly distributed in the ball in $\mathbb{R}^3$ centred at the origin with radius 1. Let D_i be the Euclidean distance from the origin of the i th point, $i = 1, \ldots, 80$, and let $\bar{D} = \frac{1}{80}\sum_{i=1}^{80} D_i$. Use the central limit theorem to find a value c satisfying $P(|\bar{D} - E[\bar{D}]| \leq c) = 0.95$.

(c)
- (i) Suppose that $\{X_n\}$ is a sequence of zero-mean random variables and X is a zero mean random variable, and suppose that $E\left[(X_n - X)^2\right] \leq C/n^p$ for every n, for some constants C and $p > 1$. Show that $X_n \to X$ almost surely.
- (ii) Suppose that $\{X_n\}$ is a sequence of non-negative random variables. Show that $E[X_n] \to 0$ as $n \to \infty$ implies that $X_n \to 0$ in probability, but that the converse is false in general.

(d) Suppose that $\{X_n\}$ and $\{Y_n\}$ are sequences of random variables and X and Y are random variables such that $X_n \to X$ in distribution and $Y_n \to Y$ in distribution. Give an example where it is not true that $X_n + Y_n$ converges to $X + Y$ in distribution. [Hint: Consider $Y = -X$]

(e) Consider the probability

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} F(X_i) - \frac{1}{2}\right| \geq \frac{1}{\sqrt{3n}}\right).$$

Use Chebyshev's inequality to bound this probability. Use the central limit theorem to approximate this probability for large n.

Problem 4. (Moment-generating function)

(a) Let $X \sim \text{Binomial}(n,p)$. Find the moment-generating function of X in terms of n and p. (Hint: X is the sum of n independent Bernoulli random variables.)

(b) Let $X_1, X_2, \ldots$ be independent and identically distributed random variables, each with a Poisson distribution with mean 1. Let $S_n = X_1 + \ldots + X_n$ for $n \geq 1$ and let $M_n(t)$ be the moment generating function of S_n.

- (i) Find the smallest n such that $P(M_n(S_n) > 1) \geq 0.99$ using the exact probability.
- (ii) Find the smallest n such that $P(M_n(S_n) > 1) \geq 0.99$ using the central limit theorem.

(c) Let X have a Gamma $(3,3)$ distribution. Conditional on $X = x$ let Z have a normal distribution with mean x and variance 2. Finally, let $Y = e^Z$. Find $E[Y]$ and $\text{Var}(Y)$.

(d) Let X be a random variable with Exponential(λ) distribution. Recall that the moment generating function of X is $M_X(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$.

- (i) Find $E[X^n]$, where n is any positive integer. You may use the mgf or compute $E[X^n]$ directly.
- (ii) Find $M_Y(t)$, the mgf of $Y = \ln X$, for $t > -1$.

(e) Let $X_1, \ldots, X_n$ be independent Poisson(1) random variables and let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Let $k > 1$ be given. Show that

$$P(\bar{X} \geq k) \leq \left(\frac{e^{k-1}}{k^k}\right)^n.$$

Problem 5. (Bivariate and Multivariate Distributions)

(a) Let N be the number of people in the state of Iowa. Suppose pN of these people support Hillary Clinton, and $(1-p)N$ of them support Donald Trump, for some $p \in (0,1)$. N is known (say $N = 3,000,000$) and p is unknown.

(i) Suppose that each person in Iowa randomly and independently decides, on election day, whether or not to vote, with probability 1/2 of voting and probability 1/2 of not voting. Let V_{Hillary} be the number of people who vote for Hillary and V_{Donald} be the number of people who vote for Donald. Show that

$$\mathbb{E}\left[V_{\text{Hillary}}\right] = \frac{1}{2}pN, \quad \mathbb{E}\left[V_{\text{Donald}}\right] = \frac{1}{2}(1-p)N.$$

What are the standard deviations of V_{Hillary} and V_{Donald}, in terms of p and N ? Explain why, when N is large, we expect the fraction of voters who vote for Hillary to be very close to p.

(ii) Now suppose there are two types of voters - "passive" and "active". Each passive voter votes on election day with probability 1/4 and doesn't vote with probability 3/4, while each active voter votes with probability 3/4 and doesn't vote with probability 1/4. Suppose that a fraction q_H of the people who support Hillary are passive and $1-q_H$ are active, and a fraction q_D of the people who support Donald are passive and $1-q_D$ are active. Show that

$$\mathbb{E}\left[V_{\text{Hillary}}\right] = \frac{1}{4}q_H pN + \frac{3}{4}\left(1-q_H\right)pN, \quad \mathbb{E}\left[V_{\text{Donald}}\right] = \frac{1}{4}q_D(1-p)N + \frac{3}{4}\left(1-q_D\right)(1-p)N.$$

What are the standard deviations of V_{Hillary} and V_{Donald} in terms of p, N, q_H, and q_D? If we estimate p by $\hat{p}$ using a simple random sample of $n = 1000$ people from Iowa. Explain why $\hat{p}$ might not be a good estimate of the fraction of voters who will vote for Hillary.

(iii) We do not know q_H and q_D. However, suppose that in our simple random sample, we can observe whether each person is passive or active, in addition to asking them whether they support Hillary or Donald. Suggest estimators $\hat{V}_{\text{Hillary}}$ and $\hat{V}_{\text{Donald}}$ for $\mathbb{E}\left[V_{\text{Hillary}}\right]$ and $\mathbb{E}\left[V_{\text{Donald}}\right]$ using this additional information. Show, for your estimators, that

$$\mathbb{E}\left[\hat{V}_{\text{Hillary}}\right] = \frac{1}{4}q_H pN + \frac{3}{4}\left(1-q_H\right)pN, \quad \mathbb{E}\left[\hat{V}_{\text{Donald}}\right] = \frac{1}{4}q_D(1-p)N + \frac{3}{4}\left(1-q_D\right)(1-p)N.$$

(b) Let X and Y be continuous random variables with joint density function

$$f(x,y) = \begin{cases} \frac{y^3}{2}e^{-y(x+1)} & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the marginal pdf of X.

(ii) Find the marginal pdf of Y and $E[Y]$.

(c) **Existence of multivariate normal.**

(i) Suppose $(X_1, \ldots, X_k) \sim \mathcal{N}(0, \Sigma)$ for a covariance matrix $\Sigma \in \mathbb{R}^{k \times k}$. Let $Y_1, \ldots, Y_m$ be linear combinations of $X_1, \ldots, X_k$, given by

$$Y_j = a_{j1}X_1 + \ldots + a_{jk}X_k$$

for each $j = 1, \ldots, m$ and some constants $a_{j1}, \ldots, a_{jk} \in \mathbb{R}$. Consider the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix}.$$

By computing the means, variances, and covariances of $Y_1, \ldots, Y_m$, show that

$$(Y_1, \ldots, Y_m) \sim \mathcal{N}\left(0, A\Sigma A^T\right).$$

(ii) Let $A \in \mathbb{R}^{k \times k}$ be any matrix, let $\Sigma = AA^T$, and let $\mu \in \mathbb{R}^k$ be any vector. Show that there exist random variables $Y_1, \ldots, Y_k$ such that $(Y_1, \ldots, Y_k) \sim \mathcal{N}(\mu, \Sigma)$. (Hint: Let $X_1, \ldots, X_k \overset{IID}{\sim} \mathcal{N}(0,1)$, and let each Y_j be a certain linear combination of $X_1, \ldots, X_k$ plus a certain constant.)

(d) (i) Let (X, Y) be a random point uniformly distributed on the unit disk $\{(x, y) : x^2 + y^2 \leq 1\}$. Show that $\mathrm{Cov}[X, Y] = 0$, but that X and Y are not independent.

(ii) Let $X, Y \sim \mathcal{N}(0, 1)$ be independent. Compute $\mathbb{P}[X + Y > 0 \mid X > 0]$. (Hint: Use the rotational symmetry of the bivariate normal PDF.)

(e) (i) A family of Yellow-Faced (YF) gophers consisting of 2 parents and 3 children are kept in a laboratory. In addition to these a family of YF gophers with 2 parents and 4 children, a family of Big Pocket (BP) gophers with 2 parents and 5 children, and a family of BP gophers with 1 mother and 4 children are also kept in the laboratory. A sample of 4 gophers is selected at random from among all the gophers in the laboratory. What is the probability that the sample consists of one adult female, one adult male, and 2 children, with both adults of the same genus (either both YF or both BP).

(ii) Let X and Y be arbitrary random variables, and let $g(\cdot)$ and $h(\cdot)$ be arbitrary real valued functions defined on $\mathbb{R}$. For each of the following statements say whether it is TRUE or FALSE. If TRUE prove it and if FALSE give a counterexample.

(I) If X and Y are uncorrelated then so are $g(X)$ and $h(Y)$ for any g and h.

(II) If $g(X)$ and $h(Y)$ are uncorrelated for all g and h then X and Y are uncorrelated.

The distribution function, $\Phi(z)$, of a standard normal random variable Note: $\Phi(-z) = 1 - \Phi(z)$ for any $z \in \mathbb{R}$.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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	0.7704	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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Useful Formulas:

- The volume of a sphere of radius r is $\frac{4}{3}\pi r^3$.
- A Maclaurin series is a Taylor series expansion of a function about 0,

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots.$$

- The generic Chernoff bound for a random variable X is attained by applying Markov's inequality to e^{tX}. This gives a bound in terms of the moment-generating function of X. For every $t \geq 0$:

$$\Pr(X \geq a) = \Pr\left(e^{t\cdot X} \geq e^{t\cdot a}\right) \leq \frac{\mathrm{E}\left[e^{t\cdot X}\right]}{e^{t\cdot a}}.$$

- Uniform distribution on the interval $(0,1)$ has pdf

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

- The standard normal distribution has pdf

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{for } x \in \mathbb{R}$$

- Beta density with parameters α and β :

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$\alpha = 1$ and $\beta = 1$ gives the Uniform distribution on $(0,1)$.

- Gamma distribution with parameters $r > 0$ and $\lambda > 0$:

$$f(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \quad E[X] = \frac{r}{\lambda}, \mathrm{Var}(X) = \frac{r}{\lambda^2}$$

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \quad \text{for } t < \lambda.$$

$r = 1$ gives the Exponential distribution with mean $1/\lambda$.

- Exponential distribution with parameter $\lambda > 0$:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \quad E[X] = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

- Normal distribution with mean μ and variance $\sigma^2 > 0$:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} & \text{if } x \in \mathbb{R} \\ 0 & \text{otherwise.} \end{cases} \quad E[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2$$

$M_X(t) = e^{\mu t + t^2\sigma^2/2}$ for $t \in \mathbb{R}$. $\mu = 0$ and $\sigma^2 = 1$ gives the standard normal distribution.

- Poisson distribution with mean $\lambda > 0$:

$$f(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & \text{if } k = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \quad E[X] = \lambda, \quad \mathrm{Var}(X) = \lambda$$

$$M_X(t) = \exp\left\{\lambda\left(e^t - 1\right)\right\} \quad \text{for } t \in \mathbb{R}.$$