

# 시계열 분석에 의한 인천공항 허브화 사업 개선을 위한 국제선 환승 여객 수 예측

20171672 수학과 이현아

## 1. 서론

올해 여러가지 이유로 인천 공항을 다섯번이나 이용하게 되었다. 그때마다 다른 나라의 공항과 비교해보았을 때 인천 공항이 세계적으로 훌륭한 평가를 받고 있는 이유를 체감할 수 있었다. 문득 인천 공항의 어떤 부분들을 높게 평가했는지 호기심이 생겨서 홈페이지의 수상 경력을 확인해보았다. 그 중에서도 '환승 서비스' 분야에서 대상을 받은 부분이 가장 인상깊었다. 공항의 능력을 평가할 때 '환승'을 의미 있게 생각한다는 것이 새로웠다. 좀 더 찾아보니 환승률은 공항의 허브 경쟁력을 평가하는 지표로 활용되고 있음을 알게 되었다. 허브 경쟁력을 인정받은 공항은 허브 공항이라 불리게 된다. 이때 허브 공항이란 지역의 중심이자 항공기들이 모이는 중심공항을 의미하고 수송량의 증가로 인한 단위당 비용을 절감할 뿐만 아니라 운항편수를 증가시킬 수 있어서 많은 나라에서 허브공항 건설을 추구하고 있다고 한다. 물론 인천 공항 또한 허브 공항으로 자리매김 하기위해 오래전부터 노력해왔다고 한다. 그리고 최근 인천국제공항공사는 4 단계 건설사업을 통해 인천공항의 인프라를 확장하여 허브화에 박차를 가하고 있다는 소식을 접했다. 하지만 인천 공항의 환승률은 최고치를 찍었던 2013 년 이후로 계속 하락세를 보이고 있으며 보통 허브 공항의 경우 환승률이 20~25% 이상인 반면 인천 공항의 경우 11.7%로 터무니없이 낮은 환승률을 보이고 있다.

이 시점에서 앞으로의 인천 공항 환승률의 추세 확인을 통한 기존의 허브화 사업들에 대한 평가가 필요하다고 생각했다.

## 2. 연구주제 및 목적

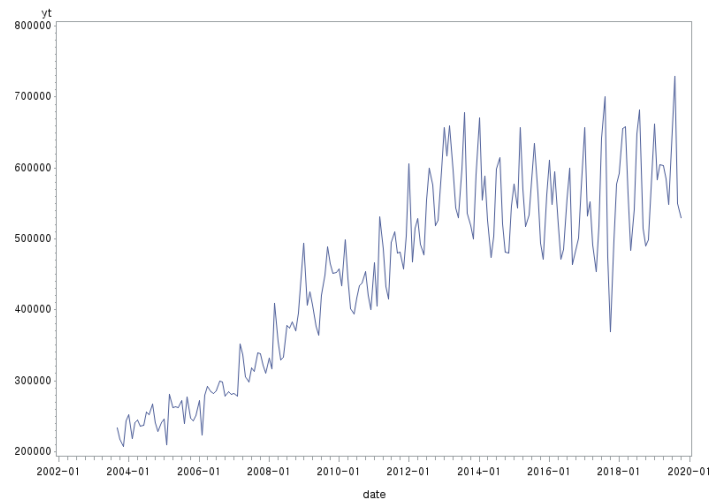
인천 공항이 동북아시아 허브 공항으로서 지위를 굳건히 할 수 있게끔 기존의 사업들에 대한 개선 필요성 여부를 판단하고자 한다. 시계열 분석을 이용하여 과거의 환승 여객 수의 자료를 통해 미래의 환승 여객 수를 예측할 수 있는 ARIMA 모형을 세우고 예측값을 구해 여객 수의 추세를 파악하고 그에 따른 평가와 대안을 제시할 것이다.

## 3. 분석 자료 설명

분석에는 인천국제공항공사의 '항공통계' 자료 중 국가통계포털 (KOSIS)에서 제공하는 <인천 공항 국제선 환승 여객 현황> 자료를 사용하였다. 그 중에서도 인천 공항을 이용하는 모든 항공사의 환승 여객 부분만을

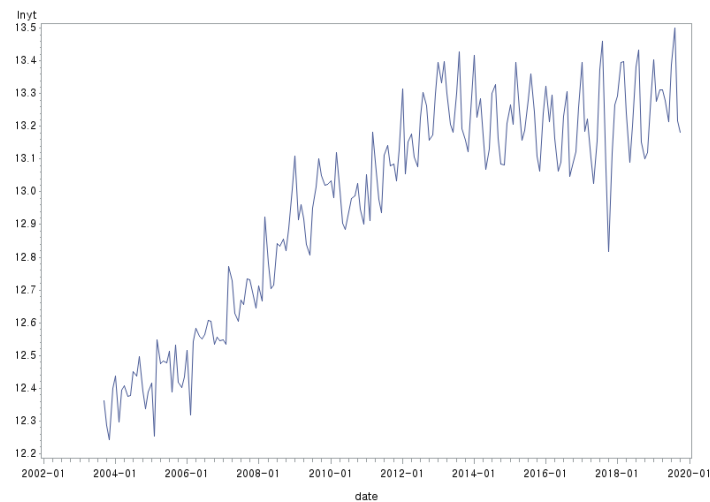
선정하였다. 이때 내항기 여객은 제외되었다. 또한 2003 년 9 월부터 2019 년 10 월 까지의 194 개 자료 분석을 통해 2019 년 11 월부터 2020 년 10 월까지의 환승 여객 수를 예측하고자 한다.

#### 4. 연구 분석



<그림 1> 인천공항 국제선 환승 여객 수( $y_t$ )의 시도표

2003 년 9 월부터 2019 년 10 월까지 인천공항 국제선 환승 여객 수에 관한 월별 시계열자료의 시도표인 <그림 1>을 보면 초기에는 시간이 경과함에 따라 수준이 증가하면서 변동 폭이 일정하지 않다가 대략 2013 년부터는 수준은 어느정도 일정해 보이나 변동의 폭만 일정하지 않은 경향을 보이고 있다. 그래서 먼저 분산안정화를 위해 로그변환을 실시하였다.



<그림 2>  $\ln y_t$ 에 대한 시도표

TRANS	LOGLIK	RMSE	AIC	SBC
NONE	-2362.42	2349477332.4	4736.84	4756.44
LOG	-2340.31	2343562821.9	4692.62	4712.23

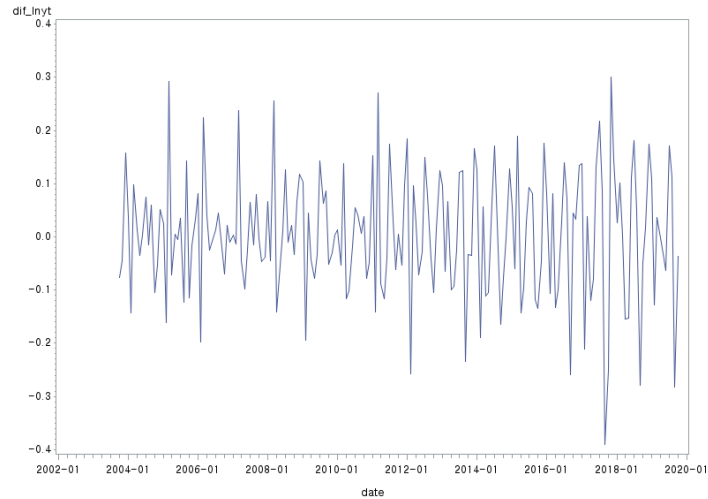
<표 1> %LOGTEST 결과

매크로 %LOGTEST의 수행 결과표(표 1)를 보면 로그변환이 더 작은 AIC와 SBC 값을 가지므로 분산안정화를 위한 로그변환은 적절한 것으로 보인다. 원 계열에 대해 로그변환을 실시한  $\ln Y_t$ 의 시도표(그림 2)를 보면 분산이 어느정도 안정적으로 되었음을 확인할 수 있다. 하지만 여전히 수준은 안정적이지 못하므로 차분을 통한 정상화가 필요할 것으로 보인다.

Mean of Working Series			12.96473		
Standard Deviation			0.327505		
Number of Observations			194		
Autocorrelations			Partial Autocorrelations		
Lag	Correlation	-198765432101234567891	Lag	Correlation	-198765432101234567891
0	1.00000		1	0.92790	
1	0.92790		2	0.04800	
2	0.86767		3	0.24085	
3	0.84440		4	0.26713	
4	0.85158		5	0.14987	
5	0.85862		6	0.08680	
6	0.85323		7	-0.01372	
7	0.83186		8	-0.10733	
8	0.79835		9	-0.15079	
9	0.76156		10	0.05148	
10	0.75280		11	0.17249	
11	0.77229		12	0.24352	
12	0.80076		13	-0.39642	
13	0.74999		14	0.01453	
14	0.70232		15	-0.00462	
15	0.68044		16	0.03220	
16	0.68456		17	-0.05584	
17	0.68385		18	-0.07945	
18	0.66804		19	0.02701	
19	0.64684		20	-0.04406	
20	0.61536		21	0.02542	
21	0.58383		22	0.02370	
22	0.57480		23	0.08815	
23	0.58993		24	0.01619	
24	0.60595		25	-0.18508	
25	0.55601		26	0.07732	
26	0.51698		27	-0.00265	
27	0.50180		28	-0.04939	
28	0.50317		29	-0.02480	
29	0.49970		30	-0.01849	
30	0.48221		31	-0.00356	
31	0.46107		32	-0.02534	
32	0.43170		33	0.05048	
33	0.40718		34	0.00935	
34	0.40041		35	-0.02314	
35	0.40819		36	0.09527	
36	0.42251				

<표 2>  $\ln Y_t$ 의 SACF와 SAPCF

<표 2>에 주어진  $\ln Y_t$ 의 SACF 와 SPACF 를 살펴보아도 SACF 가 시차가 증가함에 따라 느리게 감소하는 패턴을 보이므로 비정상성을 가지고 있다고 볼 수 있다. 따라서 앞서 급했던 것과 같이 1 차 차분을 실시하는 것이 바람직해 보인다.



<그림 3>  $\nabla \ln Y_t$ 의 시도표

$\nabla \ln Y_t$ 의 시도표(그림 3)를 통해 1 차 차분으로 수준안정화가 이루어졌음을 확인하였다.

Period(s) of Differencing		1	
Mean of Working Series		0.004241	
Standard Deviation		0.11579	
Number of Observations		193	
Observation(s) eliminated by differencing		1	
Autocorrelations		Partial Autocorrelations	
Lag	Correlation	Lag	Correlation
0	1.00000	1	-0.11773
1	-.11773	2	-0.35192
2	-.33318	3	-0.34108
3	-.20024	4	-0.27315
4	0.02286	5	-0.29199
5	0.05996	6	-0.13387
6	0.14718	7	0.00111
7	0.09911	8	0.16484
8	0.01989	9	-0.00490
9	-.23702	10	-0.16978
10	-.20216	11	-0.42300
11	-.06210	12	0.38319
12	0.65155	13	-0.00803
13	-.05495	14	0.08977
14	-.24510	15	-0.06329
15	-.19475	16	0.06387
16	0.05847	17	0.11905
17	0.09185	18	0.06041
18	0.09745	19	0.01903
19	0.05164	20	-0.02764

20	0.00031	. .		21	0.07490	.  * .	
21	-.17509	.****  .		22	-0.12953	.***  .	
22	-.18526	.****  .		23	-0.08679	.**  .	
23	-.02497	. .   . .		24	0.18388	.  ****  .	
24	0.58830	. .  *****  .		25	-0.14167	.***  .	
25	-.13667	.***  .		26	-0.01047	. .   .	
26	-.23016	.****  .		27	-0.02780	. . *  .	
27	-.12006	. . **  .		28	0.01708	. .   .	
28	0.07497	. .  * .		29	0.00232	. .   .	
29	0.07629	. .  ** .		30	-0.01393	. .   .	
30	0.07399	. .  * .		31	0.05839	. .  * .	
31	0.06437	. .  * .		32	-0.02911	. . *  .	
32	-.04269	. . . *  .		33	0.02937	. .  * .	
33	-.16401	. .***  .		34	-0.01124	. .   .	
34	-.10361	. . . **  .		35	-0.06726	. . . *  .	
35	-.03229	. . . *  .		36	0.20268	. .  ****  .	
36	0.55122	. .  *****  .		37	-0.04996	. . *  .	
37	-.16255	. .***  .		38	0.04968	. .  * .	
38	-.24005	. .****  .		39	-0.03228	. . . *  .	
39	-.10633	. . . **  .		40	0.03502	. .  * .	
40	0.11050	. . .  ** .		41	0.01886	. .   .	
41	0.08262	. . .  ** .		42	-0.01858	. .   .	
42	0.04536	. . .  * .		43	0.06204	. .  * .	
43	0.06139	. . .  * .		44	-0.00071	. .   .	
44	-.03665	. . . *  .		45	0.02240	. .   .	
45	-.18065	. .****  .		46	0.05034	. .  * .	
46	-.05204	. . . *  .		47	0.01394	. .   .	
47	-.00054	. .   .		48	-0.09394	. .**  .	
48	0.41041	. .  *****  .					

<표 3>  $\nabla \ln Y_t$ 의 SACF 와 SPACF

<표 3>에 주어진 1 차 차분한 자료에 대한 SACF 를 살펴보면 시차 12,24,45 에서  $\hat{\rho}_k$  값이 0.65155, 0.58830, 0.55122, ... 로 느리게 감소하고 있어 계절변동에 비정상성을 내포하고 있는 것으로 판단되어 계절차분을 시도하였다.

Period(s) of Differencing			1,12		
Mean of Working Series			-0.0003		
Standard Deviation			0.092295		
Number of Observations			181		
Observation(s) eliminated by differencing			13		
Autocorrelations			Partial Autocorrelations		
	Lag	Correlation	-198765432101234567891		
	0	1.00000	*****		
	1	-.18020	**** .		
	2	-.28816	***** .		
	3	0.03333	.  * .		
	4	-.07185	.  * .		
	5	-.07723	.  ** .		
	6	0.05639	.  * .		
	7	0.10309	.  ** .		
	8	0.05937	.  * .		

	Lag	Correlation	-198765432101234567891		
	1	-0.18020	**** .		
	2	-0.33139	***** .		
	3	-0.11090	.  ** .		
	4	-0.21839	**** .		
	5	-0.21360	**** .		
	6	-0.14945	*** .		
	7	-0.03811	.  * .		
	8	0.04419	.  * .		
	9	-0.09557	.  ** .		

9	-.12550	***		10	0.20393		****
10	0.16342		***	11	0.16709		***
11	0.03932		*	12	-0.32230	*****	
12	-.44997	*****		13	0.06476		*
13	0.19014		****	14	-0.12390		**
14	0.09502		**	15	0.01021		
15	-.01987			16	-0.06731		*
16	0.03017		*	17	-0.04051		*
17	0.01907			18	-0.08580		**
18	-.07325		*	19	-0.02223		
19	-.03790		*	20	-0.02579		*
20	-.00684			21	-0.08669		**
21	0.06236		*	22	0.12431		**
22	-.00289			23	0.12815		***
23	0.06812		*	24	-0.28444	*****	
24	-.09628		**	25	0.00386		
25	-.07414		*	26	-0.12001		**
26	0.04615		*	27	0.00913		
27	0.04308		*	28	-0.10218		**
28	-.02912		*	29	-0.05226		*
29	0.00941			30	0.01014		
30	0.06995		*	31	-0.06425		*
31	-.06464		*	32	-0.01949		
32	-.04191		*	33	-0.02729		*
33	0.06744		*	34	0.01697		
34	-.05194		*	35	0.02765		*
35	-.05582		*	36	0.05295		*
36	0.23742		*****	37	-0.03318		*
37	-.08642		**	38	-0.01928		
38	-.07348		*	39	-0.05459		*
39	-.04573		*	40	-0.08191		**
40	0.03303		*	41	0.02878		*
41	0.04749		*	42	-0.08507		**
42	-.06689		*	43	-0.09462		**
43	0.04076		*	44	0.07460		*
44	0.11429		**	45	-0.02860		*
45	-.12202		**	46	0.06528		*
46	0.04663		*	47	0.07247		*
47	0.03212		*	48	-0.06667		*
48	-.19498		****				

<표 4>  $\nabla_{12}\nabla\ln Y_t$ 의 SACF와 SPACF

<표 4>에 주어진  $\nabla_{12}\nabla\ln Y_t$ 의 시차 12 이전의 SACF와 SPACF는 각각 시차 2에서 절단, 시차가 증가함에 따라 지수적으로 감소하는 패턴을 보이고 있어 일반 MA(2)라고 볼 수 있다. 다음으로 12의 배수에 해당하는 시차만을 살펴보면 SACF는 시차를 따라 감소하고 SPACF는 시차 24에서 절단인 패턴을 보이므로 계절 AR(2)<sub>12</sub>라고 할 수 있다. 따라서 ARIMA(0,1,2) × (2,1,0)<sub>12</sub> 모형이 적합한 것으로 보인다. 반면에 시차 12 이전의 SACF와 SPACF를 다시 살펴보면 SACF가 시차 1 이후에 지수적으로 감소하고 SPACF도 지수적으로 감소하는 것으로 보면 일반 ARMA(1,1)이라고도 할 수 있다. 따라서 ARIMA(1,1,1) × (2,1,0)<sub>12</sub> 모형도 가능한 것으로 보인다.

✓  $ARIMA(0,1,2) \times (2,1,0)_{12}$  모형 추정

**<표 5> ML 추정결과**

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	4.92	2	0.0855	0.016	-0.005	-0.056	-0.060	-0.035	0.134
12	8.37	8	0.3979	0.042	0.049	-0.089	0.060	0.047	0.001
18	10.14	14	0.7521	0.087	-0.015	-0.024	-0.023	-0.004	-0.002
24	11.51	20	0.9319	-0.057	0.016	0.004	-0.003	0.024	-0.050
30	14.45	26	0.9666	-0.089	-0.047	-0.008	0.010	0.029	0.050
36	17.08	32	0.9856	-0.080	-0.034	0.003	-0.049	0.026	0.036
42	25.69	38	0.9361	-0.122	-0.065	-0.089	0.006	0.095	0.031
48	31.24	44	0.9260	0.031	0.015	-0.134	-0.026	0.052	-0.016

**<표 6> 포트맨토우 검정**

<표 7> 잔차의 SACF 와 SPACF

7

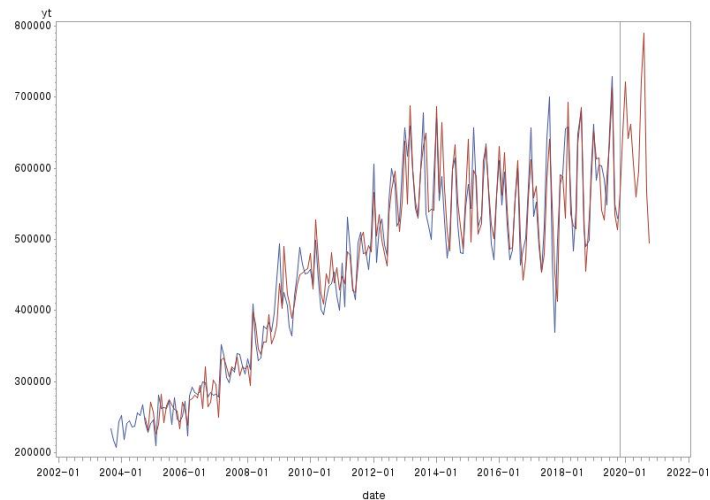
$$(1 + 0.66806 B^{12} + 0.45176 B^{24})(1 - B^{12})(1 - B) \ln Y_t = (1 - 0.26868 B^1 - 0.40306 B^2) a_t$$

(1)

모형 (1)에 근거한 향후 1 년간  $\ln Y_t$ 의 예측값(Forecast)와  $Y_t$ 의 예측값(trans-fore1)은 <표 8>과 같으며  $Y_t$ 와  $Y_t$ 의 예측값의 시계열 그림은 <그림 4>에 나타나 있다.

Forecasts for variable lnyt						
Obs	Forecast	Std Error	95% Confidence Limits		date	trans_fore1
195	13.2467	0.0657	13.1180	13.3754	2019-11	566192.80
196	13.3846	0.0813	13.2251	13.5440	2019-12	649898.27
197	13.4890	0.0842	13.3240	13.6539	2020-01	721420.32
198	13.3733	0.0869	13.2030	13.5436	2020-02	642614.65
199	13.4036	0.0895	13.2282	13.5791	2020-03	662398.37
200	13.3157	0.0921	13.1352	13.4961	2020-04	606635.78
201	13.2365	0.0946	13.0511	13.4218	2020-05	560436.80
202	13.2987	0.0970	13.1086	13.4887	2020-06	596394.13
203	13.4929	0.0993	13.2982	13.6876	2020-07	724239.52
204	13.5801	0.1017	13.3808	13.7793	2020-08	790211.69
205	13.2494	0.1039	13.0457	13.4531	2020-09	567741.36
206	13.1131	0.1061	12.9050	13.3211	2020-10	495368.26

<표 8> 모형 (1)에 근거한 향후 1 년간 여객 수 예측값



<그림 4> 원 계열과 예측값의 시도표

✓  $ARIMA(1,1,1) \times (2,1,0)_{12}$  모형 추정

Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr >  t	Lag	Variance Estimate	0.004548
MA1,1	0.84326	0.06758	12.48	<.0001	1	Std Error Estimate	0.067437
AR1,1	0.47947	0.10671	4.49	<.0001	1	AIC	-450.058
AR2,1	-0.65322	0.07139	-9.15	<.0001	12	SBC	-437.264
AR2,2	-0.43604	0.07541	-5.78	<.0001	24	Number of Residuals	181



<표 9> ML 추정 결과

<표 9>의 최우추정 결과  $\hat{\theta}_1 = 0.84326(0.06758)$ ,  $\hat{\phi}_1 = 0.47947(0.10671)$ ,  $\hat{\Phi}_1 = -0.65322(0.07139)$ ,  $\hat{\Phi}_2 = -0.43604(0.07541)$  로 모두 통계적으로 유의하다.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	12.28	2	0.0022	0.091	-0.177	-0.024	-0.062	-0.008	0.148
12	16.74	8	0.0329	0.091	0.023	-0.088	0.060	0.055	-0.009
18	18.63	14	0.1797	0.088	0.000	-0.038	-0.015	0.012	-0.007
24	20.13	20	0.4500	-0.058	0.012	0.010	0.002	0.033	-0.051
30	23.48	26	0.6058	-0.099	-0.046	-0.002	0.003	0.041	0.045
36	27.22	32	0.7071	-0.088	-0.043	0.006	-0.053	0.040	0.053
42	36.13	38	0.5563	-0.111	-0.087	-0.088	0.009	0.100	0.025
48	43.25	44	0.5036	0.039	0.011	-0.148	-0.027	0.069	-0.018

<표 10> 포트맨토우 검정

Autocorrelation Plot of Residuals			Partial Autocorrelations		
Lag	Correlation		Lag	Correlation	
0	1.00000	[-198765432101234567891]	1	0.09133	[.  **  .]
1	0.09133	[.  **  .]	2	-0.18678	[****  .]
2	-0.17688	[****  .]	3	0.01306	[.   .]
3	-0.02373	[.   .]	4	-0.09798	[. **  .]
4	-0.06190	[. *  .]	5	0.00787	[.   .]
5	-0.00776	[.   .]	6	0.12456	[.  **  .]
6	0.14757	[.  ***  .]	7	0.06467	[.  *  .]
7	0.09118	[.  **  .]	8	0.05649	[.  *  .]
8	0.02271	[.   .]	9	-0.07279	[. *  .]
9	-0.08753	[. **  .]	10	0.11682	[.  **  .]
10	0.05995	[.  *  .]	11	0.01925	[.   .]
11	0.05509	[.  *  .]	12	0.00394	[.   .]
12	-0.00906	[.   .]			

<표 11> 잔차의 SACF와 SPACF

<표 10>과 <표 11>에 주어진 모형 검진 결과  $Pr > ChiSq$  값이 0.05 보다 작은 값이 존재하고 잔차의 SACF와 SPACF 모두 시차 2에서 유의한 값이 있기 때문에  $a_t$ 가 백색잡음과정을 따른다는 가정을 만족하지 않는다. 따라서  $ARIMA(1,1,1) \times (2,1,0)_{12}$  모형은 적절하지 않다.

따라서 인천공항 국제선 환승 여객 수에 대한 최적모형으로는  $ARIMA(0,1,2) \times (2,1,0)_{12}$ 만 가능하다고 볼 수 있다.

## 5. 결론

<표 8>과 <그림 4>를 보면 2020년 1월까지는 증가하는 추세를 보이다가 다시 점차 하락하고 2020년 7월에 다시 급격히 증가했다가 2020년 10월에는 크게 감소할 것으로 예측되고 있다. 7월 즈음에 크게 증가하는 것은 여름 휴가 또는 학생들의 방학으로 인한 성수기로, 시기적인 특징이 주요인으로 보인다. 어떻게 보면 당연히 성수기에는 환승률이 증가하고 비수기에는 환승률이 크게 감소하는 모습을 보이고 있다고 할 수 있다. 이대로 비수기에 조치를 취하지 않으면 성수기와 비수기의 차이가 크게 벌어지다가

결국엔 성수기 마저도 환승률이 하락세를 보일지도 모르는 상황이다. 따라서 비수기인 달에 환승객을 유치하기 위한 노력이 특히 필요해 보인다. 그리고 지난 2017 년 대한 항공이 실행했던 중·단거리 노선 강화를 인천 공항을 이용하는 항공사들이 자발적으로 실행할 수 있게 기존보다 높은 인센티브를 제공하여 항공사들을 독려하는 것도 하나의 방법이라고 생각된다. 마지막으로 인천 공항에서의 경험으로 이야기해보자면 당시 새벽에 출발하는 비행기를 타야해서 공항에서 밤을 보낸적이 있다. 공항에 있는 캡슐 호텔이나 호텔 또는 사우나를 이용하려 하였으나 전부 만석이었고 근처에 숙박업소를 찾다가 결국은 출국 게이트 앞 의자에서 불편한 상태로 잠깐 눈을 붙였다. 이러한 경험을 했던 사람은 한 둘이 아니었다. 이용 고객을 수용할 수 있는 숙박 시설이 부족한데 환승객도 상황이 다르지 않을 것으로 보인다. 게다가 특히 환승객들은 인천 공항에 도착하는 시간이 새벽이 될지 연제가 될지는 모르는 상황이기 때문에 충분한 숙박 시설과 휴식 시설 또는 길게 체류하는 환승객들을 위한 관광 프로그램까지 마련해야할 필요가 있다고 본다.

본 연구에서 제시한 방법들 외에도 다양한 노력들을 통해 환승률을 회복하여 인천 공항이 성공적으로 동북아시아의 허브 공항으로 거듭나길 바란다.

## 참고자료

인천국제공항공사 [https://www.airport.kr/co\\_cnt/ko/cyberpr/awards/yea019/yea019.do](https://www.airport.kr/co_cnt/ko/cyberpr/awards/yea019/yea019.do)

KOSIS, 국제선 환승여객 현황(인천)

[http://kosis.kr/statisticsList/statisticsListIndex.do?menuId=M\\_01\\_01&vwcd=MT\\_ZTITLE&parmTabId=M\\_01\\_01#SelectStatsBoxDiv](http://kosis.kr/statisticsList/statisticsListIndex.do?menuId=M_01_01&vwcd=MT_ZTITLE&parmTabId=M_01_01#SelectStatsBoxDiv)

한국경제, "인천공항, 국제선 이용객 '세계 5 위' 도약"

<https://www.hankyung.com/society/article/2019021956511>

인천뉴스, "인천공항, 동북아 허브공항 맞나? 환승률 낮아지고 비행공수익 늘어나"

<http://www.incheonnews.com/news/articleView.html?idxno=119543>

머니투데이, "20 살 청년된 인천공항, "새로운 허브화 전략이 필요한 시점""

<https://news.mt.co.kr/mtview.php?no=2019013012472055601>

시사저널 e, "인천공항 환승객 유치 기대... 중·단거리 노선 강화 나선 대한항공"

<http://www.sisajournal-e.com/news/articleView.html?idxno=177580>

## 부록

```
data incheon;
infile 'C:\Users\hyunah\Desktop\timeseries\project\dp.txt';
input yt;
lnyt=log(yt);
date= intnx('month', '01sep2003'd, _n_-1);
format date yymmdd.;
run;

data incheon0;
set incheon;
if _n_ < 195;
dif_lnyt=dif(lnyt);
run;

goptions i=join;
proc gplot data=incheon0;
plot yt*date=1;
plot lnyt*date=1;
plot dif_lnyt*date=1;
run;

%logtest(incheon0,yt,out=trans,print=yes)

proc arima data=incheon0;
identify var=yt nlag=60;
identify var=lnyt nlag=60;
identify var=lnyt(1) nlag=60;
identify var=lnyt(1,12) nlag=60;
estimate p=(12,24) q=2 noint ml plot; /*1번모형*/
forecast lead=12 out=result1;
estimate p=(1)(12,24) q=1 noint ml plot; /* 2번모형*/
forecast lead=12 out=result2;
run;

/* ARIMA(0,1,2)(2,1,0) */
data result1;
merge result1 incheon;
date=intnx('month', '01sep2003'd, _n_-1);
format date yymmdd.;

trans_fore1=exp(forecast);
resid_p1=yt-trans_fore1;
run;

proc print data=result1 noobs;
var date trans_fore1 resid_p1;
where '01nov2019'd<=date<='01oct2020'd;
run;

data result1_1;
merge incheon0 result1;
by date;
if date>='01nov2019'd then yt=.;
run;

proc gplot data=result1_1;
symbol1 v=none i=join line=1;
symbol2 v=dot i=join line=33 h=0.8;
plot yt*date=1 trans_fore1*date=2/overlay href='01nov2019'd;
run;
```